

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: September 23, 2002, 21:12:04 , Search time 64.71 Seconds
(without alignments)
1208 406 Million cell updates/sec

Title: US-09-733-764-2

Perfect score: 3868

Sequence: 1 MDSIASLVICGVSLISGIV

MHFAIHHYHCKSLISLSPK 704

Scoring table:

BLOSUM62

Gap: 10 0 0 Gap: ext 0 5

Searched 747574 seqs, 11167396 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database			
A_GenSeq_032802.*			
1:	/SIDSI/qcdata/geneseq/geneseq-emb1/AA1980.DAT.*		
2:	/SIDSI/qcdata/geneseq/geneseq-emb1/AA1981.DAT.*		
3:	/SIDSI/qcdata/geneseq/geneseq-emb1/AA1982.DAT.*		
4:	/SIDSI/qcdata/geneseq/geneseq-emb1/AA1983.DAT.*		
5:	/SIDSI/qcdata/geneseq/geneseq-emb1/AA1984.DAT.*		
6:	/SIDSI/qcdata/geneseq/geneseq-emb1/AA1985.DAT.*		
7:	/SIDSI/qcdata/geneseq/geneseq-emb1/AA1986.DAT.*		
8:	/SIDSI/qcdata/geneseq/geneseq-emb1/AA1987.DAT.*		
9:	/SIDSI/qcdata/geneseq/geneseq-emb1/AA1988.DAT.*		
10:	/SIDSI/qcdata/geneseq/geneseq-emb1/AA1989.DAT.*		
11:	/SIDSI/qcdata/geneseq/geneseq-emb1/AA1990.DAT.*		
12:	/SIDSI/qcdata/geneseq/geneseq-emb1/AA1991.DAT.*		
13:	/SIDSI/qcdata/geneseq/geneseq-emb1/AA1992.DAT.*		
14:	/SIDSI/qcdata/geneseq/geneseq-emb1/AA1993.DAT.*		
15:	/SIDSI/qcdata/geneseq/geneseq-emb1/AA1994.DAT.*		
16:	/SIDSI/qcdata/geneseq/geneseq-emb1/AA1995.DAT.*		
17:	/SIDSI/qcdata/geneseq/geneseq-emb1/AA1996.DAT.*		
18:	/SIDSI/qcdata/geneseq/geneseq-emb1/AA1997.DAT.*		
19:	/SIDSI/qcdata/geneseq/geneseq-emb1/AA1998.DAT.*		
20:	/SIDSI/qcdata/geneseq/geneseq-emb1/AA1999.DAT.*		
21:	/SIDSI/qcdata/geneseq/geneseq-emb1/AA2000.DAT.*		
22:	/SIDSI/qcdata/geneseq/geneseq-emb1/AA2001.DAT.*		

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	length DB	ID	Description
1	3868	100.0	704	22	AAH31165
2	3721.5	96.0	477	22	AAH31164
3	2623	67.8	1124	15	AAH45440
4	2623	67.8	1124	16	AAH73953
5	2623	67.8	1124	23	AAH30318
6	2623	67.8	1124	22	AAH65945
7	2432.5	62.9	1124	16	AAH73951
8	2432.5	62.9	1124	16	AAH67391
9	2437.5	62.8	1124	21	AAH59046
10	2427.5	62.8	1124	21	AAH59048
11	2413.5	62.4	1127	15	AAH46427

12	2412	62.4	1117	15	AAH53146
13	2388.5	61.8	1101	20	AAH84160
14	1401.5	36.2	764	21	AAH68949
15	1354.5	35.0	963	19	AAH70540
16	1304	33.7	452	22	AAH97591
17	1304	33.7	557	22	AAH97590
18	1300	33.6	462	22	AAH97592
19	1300	33.6	518	15	AAH51003
20	1300	33.6	518	22	AAH51001
21	1300	33.6	518	22	AAH50580
22	1300	33.6	567	22	AAH97593
23	1300	33.6	567	22	AAH97597
24	1293.5	33.4	542	22	AAH13058
25	1290	33.4	497	21	AAH97172
26	1290	33.4	525	21	AAH97171
27	1290	33.4	622	21	AAH97170
28	1289	33.3	494	22	AAH13051
29	1289	33.3	528	22	AAH13061
30	1287.5	33.3	465	22	AAH13053
31	1285.5	33.2	631	10	AAH91009
32	1285.5	33.2	631	21	AAH19598
33	1285.5	33.2	631	21	AAH51079
34	1285.5	33.2	631	21	AAH59169
35	1284.5	33.2	497	22	AAH73217
36	1284.5	33.2	407	22	AAH88899
37	1283	33.2	533	22	AAH13052
38	1282.5	33.2	413	22	AAH72918
39	1282.5	33.2	413	22	AAH80900
40	1282	33.1	477	20	AAH90307
41	1281	33.1	485	13	AAH24016
42	1279.5	33.1	535	22	AAH13056
43	1278.5	33.1	396	18	AAH18574
44	1278.5	33.1	496	18	AAH18575
45	1278.5	33.1	490	21	AAH51523

ALIGNMENTS

RESULT 1

AAH31165

10 AAH31165 standard; protein; 704 AA.

XX AAH31165;

XX 02-APP-2001 (first entry)

XX Amino acid sequence of a Tek/Fe fusion protein.

XX Fusion protein, receptor tyrosine kinase; Tek, FC portion;

XX immunoglobulin-like, 140k, angiogenesis, tumour, ocular neovascularisation,
inflammatory disease; arthritis; rheumatism; psoriasis; eye disorder;
diabetic retinopathy; retinopathy; neovascular glaucoma; retinoblastoma;
retrolental fibroplasia, rubeosis; uveitis; macular degeneration;
graft neovascularisation; cancer; metastatic sarcoma; carcinoma;
wound granulation.

XX Synthetic.

XX Homo sapiens.

XX Key location/Qualifiers

XX Peptide 1..18

XX Protein /note "Signal peptide"

XX Protein /note "Tek extracellular domain fragment"

XX Protein 19..472

XX Protein 473..704

XX Protein /note "FC portion"

XX W0200075323-A1.

XX 14 DEC 2000.

XX 07 JUN 2000, 2309W0-0215706.

XX 07-JUN-1999; 990S-0137889.
 XX (IMMV) IMMUNEX CORP.
 XX Carretti DP, Borges LG, Fanslow WC;
 XX WPI; 2001-112149/12.
 XX New Tek polypeptides antagonist having a fragment of the Tek
 PT extracellular domain, useful for treating mammals with a disease
 PT mediated by angiogenesis, e.g. tumors, ocular neovascularisation or
 PT inflammatory diseases -
 XX Claim 17; Page 37-39; 43pp; English.
 XX The present sequence represents a fusion protein comprising the
 CC extracellular domain of the human receptor tyrosine kinase tek and the
 CC Fe portion of human immunoglobulin (Ig) G1. The tek fragment lacks all
 CC or part of the region containing fibronectin type III (FNIII) motifs, and
 CC retains the ability to bind at least one tek ligand. The fusion
 CC polypeptide is an angiogenesis inhibitor, and a tek antagonist. The
 CC polypeptide or soluble tek multimer, antibodies or antibody fragments are
 CC useful for treating a mammal having a disease or condition characterized by
 CC angiogenesis, e.g. a solid tumor or disease characterized by
 CC ocular neovascularisation. In particular, the tek antagonists are useful
 CC for treating or preventing inflammatory diseases (e.g. arthritis,
 CC rheumatism or psoriasis), certain eye disorders (e.g. diabetic
 CC retinopathy, retinopathy of prematurity, neovascular glaucoma,
 CC retinoblastoma, retrolental fibroplasia, rubeosis, uveitis, macular
 CC degeneration or graft neovascularisation), cancer (e.g. metastatic
 CC sarcomas or carcinomas), or wound granulation.
 XX Sequence 704 AA;

Query Match 100.0%; Score 3868; DB 22; Length 704;
 Host Local Similarity 100.0%; Pred. No. 3,40-225;
 Matches 704; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDLSASLV:GWSVLSISGTSVEVAMDLILINSPLVSDAFTSLTGIASGWPPIITIGRD 60
 DB 1 mdsaslvicgvsllsytvqamdlilinsplvsdaftsbtclsgswrphlepitigrd 60

QY 61 FEALNNQHUPLEVTJDTREWAKKVVWKEKASKINKAYFEPSVGEALPIPTMMRPQ 120
 DB 1 fealnnqhuplevtjdtrewwakvwwkrekaskinqayicedrvrgealrirmkrmq 120

QY 121 QASFLPATITMTIVKGDVNI:SPKKVLKREDAVIYKNGSP:HSVPRHFDHLEVLPH 180
 DB 121 qastlpatitmtivkdvni:spkvlkreedaviykngslihsvprhfdhlevhlph 180

QY 181 AQPDMGVYSAPYIGNIFT:SNFTFLIVER:EAUKWPE:NML:IA:MNN:V:HEH:GEC 240
 DB 181 aqpdmgvysaryiggnift:snftflvrrecadkwpecdhictacmngvchedtgec 240

QY 241 ICPPEFMGRTEKACELITFGRTEKFGSGGCKSVFCLPDPYGCSCATGMKGLCNE 300
 DB 241 icppelfmgrteakcelitfgrtekerfsggcksvfclpdpdygcscatgmkglocne 300

QY 301 ACHPGFGYHPCKLRCSNNGMCHDFPQCT:CSPGWQGIQCFRPGTIPMTPIK:VGLPDHTE 360
 DB 301 achpgfgyhpccklrscnngmchdfpqcct:cspgwqgiqcfrrpgtipmtpi:kvglpdhte 360

QY 361 VNSGFNPICKASGWLPTNEMLTVKPDGTIVLHPKDFNHTDIESVAFTIHRLLPDSG 420
 DB 361 vnskfnpickasgwlpntneemtlvkdgtivlhpkdfrnhtdiesvaftihrlppdsq 420

QY 421 VVWCSVNTVAGSVKPEPNT:SVKVL:FKPLNAPNVTCT:HNFAV:INISSEPYF:EPKST:DKT 480
 DB 421 vvwcsvntvagsvkepnt:svkvl:fkplnapnvtct:hnefav:inissepyf:epkscdkt 480

QY 481 RHCPCPAPPELLGCP:SVFLFPHKPKNTLMI:SKTPEVT:CVVVWVDSHEDPEVKFNMYDVCVE 540

DB 481 hccppcpapelggpsvflfppkpkdtilmsrtpevtcvrvdshedpe:kfnwgyvdave 540
 QY 541 VINAKTKPREEOYNSTYRVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTIISKAKGQP 600
 DB 541 vlnaktkpreedynstyrsvsltlvldqdwlngkyckckvsnkalpapiektiskakgqp 600
 QY 601 REPQVYTLPPSRDEMIKNQVSLTCLVKGIFY:SVLAWEF:SNQ:PENNYKTIPTPVLDSKGS 660
 DB 601 repqvytlppsrdeemiknvsltclvkgify:svlawef:snq:penntyktiptpvlldsdqs 660
 QY 661 FFYLSKLTVDKSRWQGNVFCSSVMHEALHNHYTOKSISLSIPGX 704
 DB 661 fflyskltvdksrwggnvfcsvmhealhnhytqkslslspgk 704

RESULT 2
 AAB31164
 ID AAB31164 standard; protein; 977 AA.
 XX AAB31164:
 XX U2-APR-2001 (first entry)
 DE Amino acid sequence of a Tek/Fe fusion protein.
 XX
 KW Fusion protein: receptor tyrosine kinase; Tek; Fe portion;
 KW immunoglobulin G1; IgG1; angiogenesis; tumor; ocular neovascularisation;
 KW inflammatory disease; arthritis; rheumatism; psoriasis; eye disorder;
 KW diabetic retinopathy; retinopathy; neovascular glaucoma; retinoblastoma;
 KW retrolental fibroplasia; rubeosis; uveitis; macular degeneration;
 KW graft neovascularisation; cancer; metastatic sarcoma; carcinoma;
 KW wound granulation.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 FH Key location/Qualifiers
 FT Peptide 1..18
 FT Protein /note= "signal peptide"
 FT Protein 19..745
 FT Protein /note= "tek extracellular domain"
 FT Protein 746..977
 FT Protein /note= "Fe portion"
 XX W0200075323-A1.
 PN 14-DEC-2000.
 PD 07-JUN-2000; 2000WO-0515705.
 PF 07-JUN-1999; qnqs-0137889.
 XX (IMMV) IMMUNEX CORP.
 PI Carretti DP, Borges LG, Fanslow WC;
 XX WPI; 2001-112149/12.
 XX New Tek polypeptides antagonist having a fragment of the tek
 PT extracellular domain, useful for treating mammals with a disease
 PT mediated by angiogenesis, e.g. tumors, ocular neovascularisation or
 PT inflammatory diseases -
 XX Claim 41; Page 34-37; 43pp; English.
 XX The present sequence represents a fusion protein comprising the
 CC extracellular domain of the human receptor tyrosine kinase tek and the
 CC Fe portion of human immunoglobulin (Ig) G1. The tek fragment lacks all
 CC or part of the region containing fibronectin type III (FNIII) motifs, and
 CC retains the ability to bind at least one tek ligand. The fusion
 CC polypeptide is an angiogenesis inhibitor, and a tek antagonist. The
 CC polypeptide or soluble tek multimer, antibodies or antibody fragments are

CC useful for treating a mammal having a disease or condition mediated by
 CC angiogenesis, e.g. a solid tumour a condition or disease characterized by
 CC ocular neovascularisation. In particular, the Tek antagonists are useful
 CC for treating or preventing inflammatory diseases (e.g. arthritis,
 CC rheumatism or psoriasis), certain eye disorders (e.g. diabetic
 CC retinopathy, retinopathy of prematurity, neovascular glaucoma,
 CC retinoblastoma, retrolental fibroplasia, rubeosis, uveitis, macular
 CC degeneration or graft neovascularisation), cancer (e.g. metastatic
 CC sarcomas or carcinomas), or wound granulation
 XX
 SQ Sequence 977 AA;

Query Match 26.2%, Score 3721.5, DR 22; Length 977;
 Best Local Similarity 72.1%; Pred. No. 3.4e 216;
 Matches 704; Conservative 0; Mismatches 0; Indels 273; Gaps 1;
 QY 1 MDLSIASLVIGVSLISGVIFGAMDLINSLPLVSIAPITSIPTLASCPDPHPPTIGPD 60
 Db 1 mdslaslvicqvsllsgfvvgamdllinslplvsdaetsltelasswrphcpitigrd 60
 QY 61 FEALMNHQJQJDLVTVDTREWAKKVVWKKREKASKINAYCEGRVRCGATPIPTMKWRQ 120
 Db 61 fealmnhqjqlvtdvtrwakkvwwkkrekaskingayfegvrigcatirtemkmrq 120
 QY 121 QASFLPALLIMVQKGLNWNISKKVLIKEEAVIYKNSFIHSVPKREKVPVILEVHLPH 180
 Db 121 qasflpallimvqkglwnwniskkvlikeeavlyknsfihsvprhevppillevhlph 180
 QY 181 AOPDQAGVYSARYIGGNLPTISATFTRIVRRCRAQKMGPCNHLCTACMNGVCHVDTCBC 240
 Db 181 aopdaqvysaryiggnlptisatftrivrrcraqkgpcnhlctacmngvchvdtdcc 240
 QY 241 ICPGFMRSRTCEKATLHTHGRPTKERQSDGRKSYVPLECPYRSPATWKGGLQVNE 300
 Db 241 icpgfmgrtceakatlhtgrtcekerqsdgrksyvtlcpdygscatwkgqlqne 300
 QY 301 ACHPCFVGPCKTICSNNGFMCIRFQDGLSICWQGLQCEKELIKWIKIIVLPDHL 360
 Db 301 achpcfvgpckircsnngfmcirfcdglswqglqcekerlikwikkivlpdhl 360
 QY 361 VNSCKFNPICKASOWPLPTNFMPLVKPQDVTIHPKDFNHTDHFSAVPTIHRILPPDSG 420
 Db 361 vnsckfnpickasowplptnfmplvkpqdvtihpkdfnhtdhfsvalftihrlppdsq 420
 QY 421 VWGCSVNTVAQWVEFFENISVKVLEPLNAPNVILTSNHFVNISSEHYEG 472
 Db 421 vwgcsvntvawvveffenisvkvleplnapnviltshfnvnisseyfgydghpkak 480
 QY 473 ----- 472
 Db 481 llykpvahyaqwhiqvtueivllyleprteylevqlvrrqeqeqhqpvrfttas 540
 QY 473 ----- 472
 Db 541 iglpprglnlfpksqtlntlwqflfssedifvwvrrsvqksdppikvgnltsv 600
 QY 473 ----- 472
 Db 601 lnlhlpregyvrvatvntkagqwsedltawtisdilppapenikisnithssaviswt 660
 QY 473 ----- 472
 Db 661 ildgysissitrykvqgknedqhdvdklnatllqyqlkglepetayqvdltaennlgs 720
 QY 473 -----HPKSCDKTHICPPCPAPHLGGVSVPLPPPKPKDT 507
 Db 721 snpafshclvltlpsqapad:gggkcpksckdktctppcpapcllqgsvflfppkpkdt 780
 QY 508 LMSRTPEVTCVVVDVSHEDPEVKWYVDGVEVHNAKTKPREQYNSTYRVVSVLTVHL 567
 Db 781 lmsrtpevtcvvvdvshedpevkfwyvdgvevhnaaktkpreeqynstyrvvsvltvhl 840

QY 568 QLWLNCKRYCKKYSNKAIPAEIKIISKAKGQPRPQVYILPPSKRPMIKNOVSIICLVK 627
 Db 841 qdwlngkeyckkysnkalpaplektiskakqqrpeqpylppstreemlknqslclck 900
 QY 628 GFYPSGTAIVFWSNCGFENNRYKTTPPVILSDSGSFFLYSKLTVDRSKWQJQNVFSSVMHE 587
 Db 901 gfypsdiavowesngpennrykttppvidsdgsfflyskltvdrskwqjnvfscvmhe 960
 QY 588 ALHNHYTKQKSLSLSPGK 704
 Db 961 alhnhytkqkslspspk 977
 RESULT 3
 AAR45440
 ID AAR45440 standard; Protein, 1124 AA.
 XX
 AC AAP45440;
 XX 25-JUN-1994 (first entry)
 XX Human orphan receptor kinase.
 DE
 XX
 KW Lrk: ligands, antibodies, PCR, amplification.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..18
 FT Region 746..772 /note= "signal peptide"
 FT Region 211..340 /note= "transmembrane region"
 FT Region 211..340 /note= "EGF-like repeat"
 FT Misc-difference 44 /note= "Cys of immunoglobulin domain"
 FT Misc-difference 102 /note= "Cys of immunoglobulin domain"
 FT Misc-difference 19..1124 /note= "claimed fragment"
 FT Misc-difference 19..745 /note= "claimed fragment"
 XX W09400469-A.
 XX 06-JAN-1994.
 XX 25-JUN-1994; 93WO-US06093.
 XX 26-JUN-1994; 94US-0905600.
 XX (IMMV) IMMUNEX CORP.
 XX Ziegler SF;
 XX WPI: 1994-026142/03.
 DB N-PSDB; AAQ55179.
 XX
 FT DNA and protein sequences for orphan receptor tyrosine kinase -
 FT and expression vectors for produ. of recombinant protein and
 FT antibodies specific for the protein, useful in research
 PS Claim 13; Fig 1; 57pp; English.
 CC Degenerate oligonucleotide primers based on the sequence conserved
 CC in the kinase domain of all receptor tyrosine kinases was used for
 CC PCR of single stranded cDNA from human placental polyA mRNA. PCR
 CC prod. HPK-6 contained a novel sequence which was used as a probe to
 CC isolate longer fragments from a human placental cDNA library. One
 CC clone contained the entire coding region and was called the ork gene.
 CC The gene prod. shown can be used as a research tool in in vitro assays
 CC for detection of ork, its ligands or their interactions.
 CC See also AAR45441.

```

XX
SQ Sequence 1124 AA:

Query Match 67.8%; Score 2623; DB 15; Length 1124;
Best Local Similarity 77.1%; Pred. No. 6,9e-150;
Matches 501; Conservative 23; Mismatches 44; Indels 82; Gaps 7;

QY 1 MDLSASLVLCGVSLLISGTVFVGMPIILINSPLVSDAETSLTCTASQWPPHPIETIGPD 60
Db 1 mdslaslvlcqvsllsqtlsvdaetlilnslplvsdaetlctasqwrphpeltlqrd 60
QY 61 FEALMNOHQDPLEVIGDVIKWKAKKVVWKKASKINGAYFCGKVRGEAIRITMKMRQ 120
Db 61 fealmnqhdpdvlgdvtrwakkvwwkaskingayfcgvrgeairitrmkmg 120
QY 121 QASFLPATLITMTVDKGNVNISFKKVLKEEDAVIYKNGSFHISVPRHVEPDLLEVILPH 180
Db 121 qasflpatlmtvdkgdvnnsfkklvleedavikngsfhsvprhvepdllevilph 180
QY 181 AQPQAGVYSARYIGNLTSAFTBLIVRRGEAKWGPECNILCTACMNGVCHEDTGC 240
Db 181 aqpqagvysaryiqnltfalsatrlivrrgeakwqpechltctamngvchdltgc 240
QY 241 ICHPGHMGPTCKACHLHIFGTCRKSGGQKCKSVVFCILPDYGGSCATGKWLQCN 300
Db 241 icphgmhgtckachlhifgtrckersgqgcksvvfcilpdygscatgkwlqcn 300
QY 301 ACHPFIYAPDKLRISVNNGEMTDFPQALSPQWGLQCTEREGIPMTPKIVDLPDHI 360
Db 301 achpfyfpdklrcsvnngemtdfpqalspqwgllqcteregipmtpkivdpdhi 360
QY 361 VNSGRFNPICKASGWPPLTNEFTLVKHDGTVLHPKDFNHTDHSVAIFTHRIILPDSG 420
Db 361 vnsgrfnpickasgwppltnfeftlvkhdgtvlhpkdfnhtdhsvalfthrilpds 420
QY 421 VVCSVNTVAGWVKRPFNISVKVLKPLNAPNVIDTGHNFVAINISSEPFYGFHFKSCDKT 480
Db 421 vvcsvntvagwvkrpfnisvklplnapnvidtghnfvainissepfygfghfkscdk 480
QY 481 HTCPGPAPPELLGSPVLFPPKPDITLMSIKRPEVICVVDVSHEDPEKFNMYV 546
Db 481 htcpgpapellgspvlfppkpditlmsikrpevicvvdvshedpekfennmyv 546
QY 474 ---gpksskkl!-----ykpnhycawqhlqvtncvltlnyicrteyeyevqlvr 521
Db 474 ---gpksskkl!-----ykpnhycawqhlqvtncvltlnyicrteyeyevqlvr 521
QY 537 --DGVEVHNAKTKPREQYNSTYRVVSVLTVLHQDWLNGKEYKCKVSNKALPAPIEKTLIS 594
Db 537 --dgvevhnaktkpreqynstyrvvsvlvtvlhqdwlngkeykckvsnkalpapiektlis 594
QY 522 raeqqeqh-----pqpvrfrft 548
Db 522 raeqqeqh-----pqpvrfrft 548
QY 595 KAKGQPRFPQVYVTPPSK--EEMTKNQVSLTCLVKGFYPS---DIAVEWE 649
Db 595 kkgqprfpqvyytpps--eemtknqvsltclvkgfyps---diavewe 649
QY 539 asig-----lppqqlnlpksgtlnltwqplfpsseddfyveve 579
Db 539 asig-----lppqqlnlpksgtlnltwqplfpsseddfyveve 579

RESULT 5
AAT30318
ID AAY30318 standard; protein: 1124 AA.

```


XX AAY30318;
 XX 15-NOV-1999 (first entry)
 XX Amino acid sequence of TEK (also known as TIE2) protein.
 XX TEK protein; TIE2 protein; receptor tyrosine kinase; T cell response;
 XX immune response; endothelial cell; tumor associated vasculature;
 XX coagulation; thrombosis; cancer; anticancer vaccine.
 XX Homo sapiens.
 XX WO9943801-A1.
 XX 02-SEP-1999.
 XX 26-FEB-1999; 99WO-GH00583.
 XX 26-FEB-1998; 98GB-0004121.
 XX (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.
 XX Durrant LG, Hewett PW, Ramage JM, Spendlove I;
 XX WPI; 1999-540586/45.
 XX New peptides containing at least one epitope from Tek receptor
 XX tyrosine kinase, used in vaccines against cancer
 XX disclosure; Fig 1; 56pp; English.
 XX The present sequence represents the TEK protein, also known as TIE2
 XX protein. TEK is a receptor tyrosine kinase. TEK contains epitopes
 XX which bind to MHC. The presentation of TEK epitopes can also stimulate
 XX helper cell and/or cytotoxic T cell responses. The immune response is
 XX directed against endothelial cells in the tumor-associated
 XX vasculature and includes production of antibodies that bind to the
 XX cells, causing coagulation and thrombosis. The immune response is
 XX targeted to endothelial cells lining blood vessels of the tumor (these
 XX cells overexpress Tek), so damage to even a few cells will kill many
 XX tumor cells. These target cells are accessible to the immune response
 XX and problems of antigenic heterogeneity, MHC loss and resistance to
 XX apoptosis (associated with epithelial cells) are unlikely to occur in
 XX normal endothelial cells. TEK epitopes (see AAY30320-24) are used to
 XX generate antibodies, and for prevention and treatment of cancer.
 XX The peptides, and recombinant DNA constructs or viral vectors that
 XX express them, are useful as anticancer vaccines to target endothelial
 XX cells that line blood vessels of the tumor.
 XX Sequence 1124 AA;
 XX
 XX Query Match 67 89, Score 2623, DB 20; Length 1124;
 XX Best local Similarity 77 14; Prod No 6, 9e-150;
 XX Matches 501; Conservative 23; Mismatches 44; Indels 82; Caps 7;
 XX
 XX 1 MDLSLASLVGVSLLSTVFGAMDIIINSPLVSDAETSLSICASQWPPHPIITIGRD 60
 XX I mdsiasivlcqvsllisqtvvgamdliinsplvsdaetslscasqwrphcpitigrd 60
 XX
 XX 61 FEALMNHQHPLEVTQGVTRFWAKKVKRKAASKINCAVFCQKVKCAIRIKRMKMQ 120
 XX 61 fealmnhqhplevtqgvtrfwakvkkrkaskincavfcqkvkcairirkrmkmg 120
 XX
 XX 121 QASFLPATLMTVDKGNVNIISPKKVIKPNDAVYKNCSEFHSVPPHVPVHLEVLPH 180
 XX 121 qasflpatlmtvdkgvnvniisppkvikpnnavyknscsefhsvpphvpvhlevlph 180
 XX
 XX 181 AQPDAGVYSARYLGSNFTSAFTPLIVRKCEAKQWSPENHLGTAQMNWVYHEEDIEP 240
 XX 181 aqpdagvysarylgsnftsaftplivrkceakqwspenhlgtamnwvyheediep 240
 XX

QY 241 ICPPHMGRTCKKACRLHTFGRTCKKRCQSGQKSYVFTFPPVQCSGATGKGIQGNF 400
 DB 241 icpphmgrtckkacrlhtfgrtckkrcsqgqkssyvicldpdyqcsatgkqlqgne 400
 QY 301 ACHPQFYGPFA*KLKSN*NNKEM*DAKPAQL*SPQWQGIQCFEFGTPPMTPKIVDLPDHE 360
 DB 301 achpqfygpfcakelhtfgrtckkrcsqgqkssyvicldpdyqcsatgkqlqgne 360
 QY 361 VASGKFNPI*KAS*WPLPTNEEMTLVKPQGVILHKKDFNHTGHFSAVILHKKILLPQSSG 420
 DB 361 vnsqkfnpicasqwpplneemtlvkpqtvlhpkdfnhtghfshvalitihkllppdsq 420
 QY 421 VVCSVNTVAGVVEKPFNISVKVLPKPI*NAPIVITGHNPAVINISSHPYFGRPKSCDKT 480
 DB 421 vvcsvntvagvvekpfnisvkvlpkplnapovidlqhafvlnisspeylqg----- 473
 QY 481 HTCPGCPAPPELLGGPSVFI*FPKPKDILMSIKTPETVCVVVDVSHEDPEVKENWVY---- 536
 DB 474 ----gfksskll-----yknhycawghlqfnctvctgyleftrgyleqlyr 521
 QY 537 --DGEVHNAKTKPREEQYNSTYFVVSVLTLVHQLWLN*KEYKCKVSNKALPAPIKRTIS 594
 DB 522 rgeqgeqh-----lpprdalnlpksqtllnitwqplfpsseddyveve 579
 QY 595 KAKGQRPFGVYTLDPNR--EEMIKNOVSLTCLVKGPYPS---DIAYEWE 639
 DB 539 aslg-----lpprdalnlpksqtllnitwqplfpsseddyveve 579
 RESULT 6
 AAC65945
 ID AAG65945 standard; protein: 1124 AA.
 AC AAG65945;
 XX 11-FEB-2002 (first entry)
 XX Amino acid sequence of human tie-2.
 XX Tie-2: catalytic domain; receptor tyrosine kinase, crystalline; human;
 XX cytosolic; vasotropic; antianemic; antiarteriosclerotic; nephrotropic;
 XX ophthalmologic; hepatotropic; antithyroid; antiinflammatory; antiulcer;
 XX gastrointestinal; antirheumatic; osteopathic; antialthritic; hemostatic;
 XX antipsoriatic; dermatological; immunosuppressive; antibacterial.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 XX Domain 802..1124
 XX /note= "catalytic domain; specifically claimed fragment"
 XX
 XX WO200172778-A2.
 XX 04 JC; 2001.
 XX 20 MAR 2001, 2001WO-US08853.
 XX 29-MAR-2000; 2000RS 192920P.
 XX (BADI) BASF AG.
 XX Rump NJ, Arnold LD, Dixon RW, Hoeffer HW, Allen K, Hellmanna G;
 XX WPI; 2001-648437/74.
 XX Crystalline polypeptide useful for identifying inhibitors of a tie 2
 XX protein as well as determining the three dimensional structure of a
 XX polypeptide comprising the catalytic domain of a tie-2 polypeptide -
 XX claim 6; Fig 1; 42pp; English.
 XX The invention relates to a crystalline polypeptide, comprising the
 XX catalytic domain of a receptor tyrosine kinase Tie 2 protein. The

crystalline forms are useful for identifying inhibitors of a Tie-2 protein as well as determining the three dimensional structure of a the catalytic domain of a Tie-2 polypeptide. A Tie-2 inhibitor may be used to treat a Tie-2 dependent condition in a patient (especially a human), where the condition is characterized by excessive vascular proliferation e.g. a hyperproliferative disorder, cancer (e.g. sarcoma, osteoma, melanoma, lymphoma, and leukemia), a cardiovascular condition (e.g. atherosclerosis, ischemia, anemia, and vascular leakage disorders), an ocular condition (myopia, chronic retinal detachment, conjunctivitis, retinopathy, and macular degeneration), von Hippel Lindau disease, hemangioid, psoriasis, breast disease, polycystic kidney disease, fibrosis, sarcoidosis, cirrhosis, thyroiditis, Oslar-Weber-Rendu disease, chronic inflammation, synovitis, inflammatory bowel disease, Crohn's disease, rheumatoid arthritis, osteoarthritis, psoriatic arthritis, and ulcer or sepsis, especially where the disorder involves aberrant endothelial pericytobothelial interactions. The Tie-2 inhibitor may be used to decrease fertility, and promote angiogenesis or vasculogenesis (in combination with a pro-angiogenic growth factor). The present sequence represents the human Tie-2 protein sequence.

XX Sequence 1124 AA;

Query Match 67.8%; Score 2623; DB 22; Length 1124;
 Best Local Similarity 77.1%; Pred. No. 6.9e 150;
 Matches 501; Conservative 23; Mismatches 44; Indels 82; Gaps 7;

QY 1 MDLSASLVGVSLTSCIVFGAMDLILINSPLVSNAPISLTCTASQWRPHPTITGRD 60
 Db 1 mdslaslvqvslstscivfllinsplvsnapisltctasqwrphptitgrd 60

QY 61 FEALMNHQDPLEVTDVTVREWAKVWKKRASKINAYETGVRVREATPTPTMKRPQ 120
 Db 61 fealmnhqdplevtdvtrwakkvwwkkraskinayetgvrvrreaptptptmkrpq 120

QY 121 QASELPATLTMTVDKGDNNVISFKKVLKEEDAVIYKNSFIHSVPREVPDILLEVHLPH 180
 Db 121 qaselpatltmtvdkgdnnvisfkvllkeedavlyknsfihsvprevpdpillevhlph 180

QY 181 AQDQAGVVSARYIGENLTISAFIRLVKPGFAQKWPCCNHLCTACMNNVCVCHEDIGEC 240
 Db 181 aqdqagvvsaryiqgnltisatrlvrceaqkwpccnchlctacmnnqvchedigec 240

QY 241 ICPDPPMGHTTEKAEELITFSGTKEKSSQLEAKSYVEFLPDYGSGLAIGWKGLOGNE 300
 Db 241 icppdppmghttekaeelitfsgtkekssqleakksyveflpdygsglalgwkglogne 300

QY 301 ACIPGFGPGPKLRPSNNRGMTPGFAGTASPGWQSLQVPEKIPRMTPKIVLPDHPHF 360
 Db 301 achpfgpgpklrpsnnrgmtpgfagtaspgwqslqvpekprmtpkivlpdhp hf 360

QY 361 VNSGKNPIKCKASWDQDINHPMTVKVPCIVIHVQNNRPHRSVAFTPIHPLIPDSC 420
 Db 361 vnsqknpiikckaswdqdinhpmtkvpcivihvqnnrphrsvaftpihplipdsc 420

QY 421 VVWGSNTVAGVWVKEPNTSVKVLKPELNAPNVTCTGHNFAVINLSPEYGEKSCDKT 480
 Db 421 vvwgsntvagvwvkepntsvkvllkpeelnapnvtctghnfavinlspeygekscdk 480

QY 481 HTCPGPCAPPELLGSPVLEFPFKPKDITLISRTPEVTICVVDVSHEDPEVKFMVY 536
 Db 481 htcpGPCAPPELLGSPVLEFPFKPKDITLISRTPEVTICVVDVSHEDPEVKFMVY 536

QY 537 --DGVFVHNAKIKPRHGYNSIYRVVSVLIVLHQWLINGKEYCKVSNKALPAPLTKTIS 594
 Db 537 --dgvfVHNAKIKPRHGYNSIYRVVSVLIVLHQWLINGKEYCKVSNKALPAPLTKTIS 594

QY 595 KAKGAPPEPVWYTLPPSP FEMTFNIVSLTCLVKEGFYPS---DIAVEWE 639
 Db 595 kagapPPEPVWYTLPPSP fEMTFNIVSLTCLVKEGFYPS---DIAVEWE 639

QY 599 asig-----lpprgnlilpksqtlnltwqpfpseddqfyove 579
 Db 599 asig-----lpprgnlilpksqtlnltwqpfpseddqfyove 579

RESULT 7

AAR73951
 ID AAR73951 standard; Protein; 1123 AA.
 XX AC AAR73951;
 XX DT 21-JAN-1996 (first entry)
 XX DE Mouse Tie-2 receptor tyrosine kinase protein.
 XX KW Tie-2, receptor tyrosine kinase, DNA primer; cancer; angiogenesis;
 XX KW vasculogenesis.
 XX OS Mus musculus.
 XX FH Key Location/Qualifiers
 FT Cleavage site 22..23 /note- "signal peptide cleavage site"
 FT Misc-difference 44 /note- "involved in sulphydryl bonding of Iq domain"
 FT Misc-difference 102 /note- "involved in sulphydryl bonding of Iq domain"
 FT Region 211..251 /note- "epidermal growth factor-like repeat"
 FT Region 255..298 /note- "epidermal growth factor like repeat"
 FT Region 302..340 /note- "epidermal growth factor-like repeat"
 FT Domain 437..720 /note- "includes 3 fibronectin type-III domains"
 FT Domain 829..921 /note- "tyrosine-kinase domain"
 FT Domain 936..1081 /note- "tyrosine kinase domain"
 FT Region 135..137 /note- "RGD triplet"
 XX W09513387-AL.
 XX 18-MAY-1995.
 XX 12-NOV-1994; 94W0-EP03767.
 XX 12-NOV-1993; 93US-0152552.
 XX (PLAC) MAX PLANCK GES FORDERUNG WISSENSCHAFTEN.
 XX Risau W;
 XX WPI: 1995-194105/25.
 XX N-PSDB; AAQ91999.
 XX New Tie-2 receptor tyrosine kinase and related nucleic acid and methods for detecting Tie-2 modulators for treating eg cancer, associated with angiogenesis and vasculogenesis
 XX Claim 3; Page 45; 81pp; English.
 XX This protein may be expressed recombinantly in a host cell. The expressed protein may be used to treat diseases, or processes, associated with angiogenesis and vasculogenesis, or cancer. Cells that express the protein are used in screening procedures and recombinant proteins can be used for affinity purification of Tie-2 ligand.
 XX Sequence 1123 AA;

Query Match 67.8%; Score 2432.5; DB 16; Length 1123;
 Best Local Similarity 68.1%; Pred. No. 2.1e-138;
 Matches 481; Conservative 50; Mismatches 94; Indels 81; Gaps 14;

QY 1 MDLSASLVGVSLTSCIVFGAMDLILINSPLVSNAPISLTCTASQWRPHPTITGRD 60
 Db 1 mdslaslvqvslstscivfllinsplvsnapisltctasqwrphptitgrd 60

Db	1	m	d	s	a	q	a	l	v	c	q	s	i	l	y	g	v	e	a	m	d	i	l	i	n	s	i	p	r	v	s	d	a	e	t	s	l	l	c	i	a	s	w	i	n	p	h	e	p	i	l	a	r	d	60
QY	61	F	H	A	L	M	N	O	H	D	E	R	V	T	G	V	T	F	W	A	K	K	V	W	K	R	K	A	S	K	I	N	A	Y	F	C	S	E	S	V	G	F	A	I	P	I	R	K	M	K	120				
Db	61	f	o	a	l	m	n	o	h	d	e	r	v	t	g	v	t	f	w	a	k	k	v	w	k	r	k	a	s	k	i	n	a	y	f	c	s	e	s	v	g	f	a	i	p	i	r	k	m	k	120				
QY	121	Q	A	S	F	E	P	A	T	I	M	V	D	K	I	N	N	I	S	F	K	K	V	L	K	E	E	R	A	V	I	Y	K	N	S	F	T	H	S	V	P	P	H	E	V	P	I	L	E	V	H	P	180		
Db	121	q	a	s	f	e	p	a	t	i	m	v	d	k	i	n	n	i	s	f	k	k	v	l	k	e	e	r	a	v	i	y	k	n	s	f	t	h	s	v	p	p	h	e	v	p	i	l	e	v	h	p	180		
QY	181	A	Q	H	D	A	C	V	T	S	A	R	V	I	S	A	P	T	R	I	L	P	P	C	P	A	C	K	W	G	D	P	C	N	H	L	I	C	A	M	N	N	V	C	H	E	P	240							
Db	181	a	q	h	d	a	c	v	t	s	a	r	v	i	s	a	p	t	r	i	l	p	p	c	p	a	c	k	w	g	d	p	c	n	h	i	c	a	m	n	n	v	c	h	e	p	240								
QY	241	I	C	P	P	C	P	M	C	H	E	P	C	H	E	P	C	H	E	P	C	H	E	P	C	H	E	P	C	H	E	P	C	H	E	P	C	H	E	P	C	H	E	P	C	H	E	P	300						
Db	241	i	c	p	p	c	p	m	c	h	e	p	c	h	e	p	c	h	e	p	c	h	e	p	c	h	e	p	c	h	e	p	c	h	e	p	c	h	e	p	c	h	e	p	c	h	e	p	300						
QY	301	A	C	H	E	P	C	H	E	P	C	H	E	P	C	H	E	P	C	H	E	P	C	H	E	P	C	H	E	P	C	H	E	P	C	H	E	P	C	H	E	P	C	H	E	P	C	H	E	P	360				
Db	301	a	c	h	e	p	c	h	e	p	c	h	e	p	c	h	e	p	c	h	e	p	c	h	e	p	c	h	e	p	c	h	e	p	c	h	e	p	c	h	e	p	c	h	e	p	c	h	e	p	360				
QY	361	V	N	S	G	K	N	P	I	C	K	A	S	G	W	P	I	N	E	M	T	I	V	K	P	D	T	V	L	H	K	P	D	N	H	T	D	H	F	S	V	A	I	T	H	R	I	P	P	D	S	G	420		
Db	361	v	n	s	g	k	n	p	i	c	k	a	s	g	w	p	i	n	e	m	t	i	v	k	p	d	t	v	l	h	k	p	d	n	h	t	d	h	f	s	v	a	i	t	h	r	i	p	p	d	s	g	420		
QY	421	V	W	C	S	V	N	T	V	A	G	W	E	K	P	E	N	I	S	V	K	V	L	K	P	L	N	A	N	V	I	D	C	H	N	F	A	V	I	N	S	E	H	I	F	C	-----	472							
Db	421	v	w	c	s	v	n	t	v	a	g	w	e	k	p	e	n	i	s	v	k	v	l	k	p	l	n	a	n	v	i	d	c	h	n	f	a	v																	

RESULT

RESOL
AAR67391

AAR67391
 ID AAR67391 standard; Protein: 1123 AA.

XX
XX

AC AAR67391;

XX
XX

DT 24-AUG-1995 (first entry)

XX

DE Murine tyrosine kinase

XX XX

KW Murine tyrosine kinase receptor; tie

XX

OS
yy
Mus musculus.

XX
VCE

File Key
File Peptide

peptidic

1.4

PP	06-MAY-1994:	93TP-0129912.
XX	(SUDA/)	SUDA T.
XX	(YAMA)	YAMANOUCHI PHARM CO LTD.
XX		
XX	WPI:	1995-032331/05.
XX		N-P8DB: AAQ75334.
XX		
XX	UNA coding a tie-2 receptor and a tie 2 receptor - used in the	
XX	diagnosing hematosis	
XX		
PS	Claim 2;	Pages 13-17; 17pp; Japanese.
XX		
CC	AAQ75334	encodes AAR67391 the murine 'tyrosine kinase receptor tie-2.
CC	The cDNA	is thought to participate in hematosis, and can therefore
CC		be used in the diagnosis of hematosis.
XX		
XX	Sequence	1123 AA;
XX		
SQ		

Qy 610 PSRHEM-----TKNQVSLTCLVKGFYPSDIAVFWFSNQGPENNY 648
 | | : | | : | | : | | : | | : | | : | | : | | : | | :
 Db 639 qpapenikisnlttdstamvswt-ivdgysissiliirykvqgknedqh 683

```

RESULT 9
ID AAY59046
XX AC AAY59046:
XX DT 07-MAR-2000 (first entry)
XX DE Mouse receptor tyrosine kinase protein
XX KW Receptor tyrosine kinase protein; angiogenesis; cardiogenesis; mouse;
XX KW tumorigenesis.
XX OS Mus musculus.
XX PN US5998187-A.
XX PD 07-DEC-1999.
XX PF 23-APR-1997; 97US-0838957.
XX PR 20-MAR-1994; 94US-0278089.
XX PR 30-JUL-1992; 92US-0921795.
XX PR 29-APR-1994; 94US-0235408.
XX PA (MOUN ) MOUNT SINAI HOSPITAL CORP.
XX PI Yamauchi TP, Breitman J, Dumont DJ, Rossant J, Breitman ML;
XX WPI: 2000-052545/04.
XX N-PSDB: AAZ40648.
XX PT Novel receptor tyrosine kinase protein involved in angiogenesis,
XX PT cardiogenesis and tumorigenesis.
XX PS Claim 4; Fig 1; 124pp; English.
XX CC The invention provides novel receptor tyrosine kinase protein sequences
XX CC (AAY59046-48) and nucleic acids encoding the polypeptides. The isolated
XX CC and purified polypeptides are useful for studying the developmental
XX CC expression of a receptor tyrosine kinase protein which may have a role
XX CC in angiogenesis, cardiogenesis and tumorigenesis. The present sequence
XX CC represents a mouse receptor tyrosine kinase protein.
XX SQ Sequence 1118 AA:

Query Match: 62.8% Score 2427.5, DB 21, Length 1118,
Best Local Similarity 64.0% Prod. No. 4.2e-138;
Matches 480; Conservative 50; Mismatches 95; Indels 81; Gaps 14;

QY 1 MUSLASLVGVSLLSGVTSVAMDLILINSLPLVSDAETSLTCTASGWRPHEPITIGRD 60
DB 1 mdsIagIvgvsIIlyvveqamdlilinsIplvsdaetsltctasgwhphepitigrd 60
QY 61 FHALMNOHQDDPEVTGVIRWAKVWKKPKASKINCAVEFCRGVRGSAIKRKMKMQ 120
DB 61 teaimmqdplevqdvirewakvwwkrekaskinqavicegrvrgqairrtmkmq 120
QY 121 GASFLPATLTMFVKGQNVNISPKKVILKEEDAVLYKNGSFTHSVPRHEVFDILEVHLPH 180
DB 121 qusflpatlmtvdrgdnvnistkklkeedavlykngsfthsvprhevpdilevhlph 180
QY 181 AQPQDAGVYSARYTGSNLEISAFTELLIVRRLEAKWKEPEGNHLLTA:MNGVGHEDIGEC 240
DB 181 aqpqdagvysarytgrlfsaftrlvrrleakwkepegnhllta:mngvghedigec 240
QY 241 ICPPGFMCRICEKAELEHIFCKEKEKESGSGGCKSVYVPEIIPQYGVSAVWKGLEWNP 300
DB 241 icppgmtericekceplartckesapeqcksvyiclpqyqscatqawqlqene 300
QY 301 ACHUCFYCHCKLKCSCNCHMCDFPGGGLCSQGWGGICCFPRGGLPMPKIVLDPHLE 360
DB 301 achucfychcklkcscnchmcdfpgggllcsqgwggiccfprggllpmpkivldphle 360

```

```

DB 301 acpsgyyqdccklrchctneeedrfqgclcsqgwqglqcekurprmtpqicldpdhic 360
QY 361 VNSCKFNEICKASGMPLEPTNEEMTLVKPCHGTVLHPKPNHUTHPFSAVITTHILPDPDG 420
DB 361 vnsckfneickasgmpleptneemtlvkpchgTVLHPKPNHUTHPFSAVITTHILPDPDG 420
QY 421 VVCSVNTVAGMVEKPFNISVKVLKPIINAPNVIDTGHNFVAVINISSEPYFG----- 472
DB 421 vvcsvntvAGMVEKPFNISVKVLKPIINAPNVIDTGHNFVAVINISSEPYFG----- 472
QY 473 -----EPKS-----CDKTHITCPGPAPPELLGAP-----S 496
DB 473 -----EPKS-----CDKTHITCPGPAPPELLGAP-----S 496
QY 497 VFLEPPKPKDTLMISRPPEVTCVVVDVSDHEDPEVKFNWYVDGVEVHNAKTKKREEQYNS 555
DB 497 vfleppkpkdtlmisrppevTCVVVDVSDHEDPEVKFNWYVDGVEVHNAKTKKREEQYNS 555
QY 556 TYRVVSVLT-VRHQLWINGKEYCKVSNKALPAPTEKISKAKQ-PRRPPQVYV-----LP 609
DB 556 tyrvvsvlt-VRHQLWINGKEYCKVSNKALPAPTEKISKAKQ-PRRPPQVYV-----LP 609
QY 590 -ik:pqnlt:slsal:preq:trra-----efy---vevrrslqittdsqdh 648
DB 590 -ik:pqnlt:slsal:preq:trra-----efy---vevrrslqittdsqdh 648
QY 610 PSREEM TKNQVSLTCLVKGIFYPSDIAVEWESNQPENNY 648
DB 610 psreem TKNQVSLTCLVKGIFYPSDIAVEWESNQPENNY 648
QY 639 ppenikisnltstamvswf-vdgysissitlrykvvgkncdqh 683
DB 639 ppenikisnltstamvswf-vdgysissitlrykvvgkncdqh 683

RESULT 10
AAY59048
ID AAY59048 standard; Protein: 1122 AA.
XX AC AAY59048:
XX DT 07-MAR-2000 (first entry)
XX DE Mouse tek receptor tyrosine kinase protein.
XX KW Receptor tyrosine kinase protein; angiogenesis; cardiogenesis; mouse;
XX KW tumorigenesis.
XX OS Mus musculus.
XX PN US5998187-A.
XX PD 07-DEC-1999.
XX PF 23-APR-1997; 97US-0838957.
XX PR 20-JUL-1994; 94US-0278089.
XX PR 30-JUL-1992; 92US-0921795.
XX PR 29-APR-1994; 94US-0235408.
XX PA (MOUN ) MOUNT SINAI HOSPITAL CORP.
XX PI Yamauchi TP, Breitman J, Dumont DJ, Rossant J, Breitman ML;
XX WPI: 2000-052545/04.
XX N-PSDB: AAZ40640.
XX PT Novel receptor tyrosine kinase protein involved in angiogenesis,
XX PT cardiogenesis and tumorigenesis.
XX PS Claim 1; Fig 11B; 123pp; English.
XX CC The invention provides novel receptor tyrosine kinase protein sequences
XX CC (AAY59046-48) and nucleic acids encoding the polypeptides. The isolated
XX CC and purified polypeptides are useful for studying the developmental
XX CC expression of a receptor tyrosine kinase protein which may have a role
XX CC in angiogenesis, cardiogenesis and tumorigenesis. The present sequence
XX CC represents a mouse tek receptor tyrosine kinase protein.
XX SQ Sequence 1122 AA:

```

Query Match 62.8%; Score 24127.5; DB 21; Length 1122;
 Best Local Similarity 68.0%; Pred No 4 20-138;
 Matches 480; Conservative 50; Mismatches 95; Indels 81; Gaps 14;

QY 1 MDSLASVLCGVSLLSGIVEGAMDLILLNSLPLVDAFTSLTCTASGWPHPPTTIGHD 60
 DB 1 mdsilagivcgslllygvegamdlillnsplvdaftslctctasgwhphtitgrd 60

QY 61 FEALMNHQDPLEVQDVTREWAKKVVWKEKASKINAVFVEGVPVFEAIPITMKMPQ 120
 DB 61 fealmnhqdplevqdvtrwakkvwwkekaskinavfvegvpefaipitmkmpq 120

QY 121 QASFLPATLITMVDKGNVNISKKVILKPEDAVIYKNGSFTHSVPRHEVPDILKVLHPH 180
 DB 121 qasflpatlitmvdkgvnvnskkvilkpedaviyknsgfthsvprhevdpdilevhlph 180

QY 241 ICPPGEMGRTCEKACELHTIEFTKERCSEGECKSYVFLCPDPYGCSCATGKGLQCN 300
 DB 241 icppgemgrtceakacelhrtfctkeresgecksyvflcpdpypgcsctatgkglqcn 300

QY 301 ACHPGFYPDCKIKSCNNGEMCDRFGCLCSPGWOGLQCRKGIIPRMTPKIVDIPDHL 360
 DB 301 achpgfypdckikscnngemcdrfgclcspgwoqlcqrkgiiprmtpkivdipdhl 360

QY 361 VNSGFNPICKASGWPPTNEEMTLVKKPLKPLNAPNVLDTHGNEFAVINISSEPYFG 420
 DB 361 vnsghfnpickasgwpptneemtlvkkplkplnapnvldthgnefavinisseyfg 420

QY 421 VWVGSVNTVAGMVEKPFENISVKVLPKPLNAPNVLDTHGNEFAVINISSEPYFG 480
 DB 421 vwvgsvntvagmvekpfenisvkvlpkplnapnvldthgnefavinisseyfg 480

QY 473 -----EPKS-----CDKTHCPGPCPAPPELLGGP-----S 496
 DB 473 -----EPKS-----CDKTHCPGPCPAPPELLGGP-----S 496

QY 481 ILYKPVQAWKYIEVTEITLNYLEPRDYELC--VQLARPGEGEQHPQVRRITLAC 538
 DB 481 ilypvqawkyievteltnyleprdyelc--vqlarpggeqhpqpvrrittac 538

QY 497 VFLPFPKPKDTLMTLSRTP-EVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTPRREQYNS 555
 DB 497 vflpfpkpkdtlmtlsrtp-evtcvvvdvshedpevkfnwnyvdgvevhnaaktprreqyns 555

QY 539 IGLPPRGLSILPKSQALNTWQPIFNSED-----CFY----VEVCRSLQTTSDGQN- 589
 DB 539 iglpprglsilpksqalntwqpfinsed-----cfy----vevcrslqtsdgn- 589

QY 556 TYPVSVLIT-VLHQFWLWLNCKEYFKVSNKALPAPIERTISKAKGQ-PRPEQVVT----LP 609
 DB 556 typvsvlit-vlhwfwlwlnckeyfkvsnkallpapiertiskakgq-prpeqvvt----lp 609

QY 590 IKYVGNLTSVLSNLYPQGYTVRA-----RVNTKADGEWSCELRATWLSLILP 638
 DB 590 ikvgnltsvlsnlypqgytvra-----rvntkadgewseclratwlsdilp 638

QY 610 PSREEM-----PKNOVSLTCLVKCHYPSPDIAVWESNCOQPNY 648
 DB 610 psreem-----pknovsltclvkchypspdiahwesncqpny 648

QY 639 PGPENIKSLNITDSTANVSWL-IVDQYSISSIIIRYKVAQKNEDQH 683
 DB 639 ppenikslnitdstanvswl-ivdqysissiiirykvaqknedqh 683

RESULT 11
 ID AAR48627
 AC AAR48627 standard; Protein; 1122 AA.
 DT 31-JUL-1994 (first entry)
 DE Protein tyrosine-kinase.
 KW Protein-tyrosine-kinase; tek; transcriptional regulatory element;
 KW endothelium; gene therapy; vascular disease.
 OS Mus musculus.
 PN W09404694-A.
 PD 03-MAR-1994.
 XX

25-AUG-1993; 93WO-CA00352.
 25-AUG-1992; 92US-0334493.
 (MOON) MOUNT SINAI HOSPITAL CORP.
 Breitman MJ, Dumont D, Gradwohl GC;
 WPI: 1994-083208/10.
 N-PSDB; AAQ56697.
 Isolated transcriptional regulatory element - for directing
 expression of gene specifically in cells of endothelial lineage
 Disclosure; Page 42-46; 74pp; English.
 CC cDNA of sequence AAQ56697, derived from mouse strain CD-1 embryo
 heart, encodes a protein tyrosine-kinase (tek) of sequence AAR48627
 that is expressed during mouse cardiogenesis. An additional cDNA
 sequence, given in AAQ56698 and encoding protein AAR48628, was isolated
 from a phage lambda-gt10 mouse embryo cDNA library. DNA of sequence
 CC AAQ56699 was isolated from a mouse genomic phage library using tek
 cDNA as probe, and included the transcriptional regulatory element
 of tek comprising the initiation codon and untranslated sequences.
 CC The element may be used in gene therapy to introduce foreign genes
 into endothelial cells to correct/prevent vascular disease.
 XX
 SQ Sequence 1122 AA;

Query Match 62.4%; Score 2413.5; DB 15; Length 1122;
 Best Local Similarity 67.7%; Pred. No. 2.9e-137;
 Matches 478; Conservative 50; Mismatches 97; Indels 81; Gaps 14;

QY 1 MDSLASVLCGVSLLSGIVEGAMDLILLNSLPLVDAFTSLTCTASGWPHPPTTIGHD 60
 DB 1 mdsilagivcgslllygvegamdlillnsplvdaftslctctasgwhphtitgrd 60

QY 61 FEALMNHQDPLEVQDVTREWAKKVVWKEKASKINAVFVEGVPVFEAIPITMKMPQ 120
 DB 61 fealmnhqdplevqdvtrwakkvwwkekaskinavfvegvpefaipitmkmpq 120

QY 121 QASFLPATLITMVDKGNVNISKKVILKPEDAVIYKNGSFTHSVPRHEVPDILKVLHPH 180
 DB 121 qasflpatlitmvdkgvnvnskkvilkpedaviyknsgfthsvprhevdpdilevhlph 180

QY 181 AQPDLASVYSARYIIGNLFTSAFTLIVKPEAKWKEPEINHLDTA-MNNVY-HEDTGE- 240
 DB 181 aqpdagvysaryiignlftsaftrlivrceagkwgpcsrpctctckngvchcdtgc- 240

QY 241 ICPPGEMGRTCEKACELHTIEFTKERCSEGECKSYVFLCPDPYGCSCATGKGLQCN 300
 DB 241 icppgemgrtceakacelhrtfctkeresgecksyvflcpdpypgcsctatgkglqcn 300

QY 301 ACHPGFYPDCKIKSCNNGEMCDRFGCLCSPGWOGLQCRKGIIPRMTPKIVDIPDHL 360
 DB 301 achpgfypdckikscnngemcdrfgclcspgwoqlcqrkgiiprmtpkivdipdhl 360

QY 361 VNSGFNPICKASGWPPTNEEMTLVKKPLKPLNAPNVLDTHGNEFAVINISSEPYFG 420
 DB 361 vnsghfnpickasgwpptneemtlvkkplkplnapnvldthgnefavinisseyfg 420

QY 421 VWVGSVNTVAGMVEKPFENISVKVLPKPLNAPNVLDTHGNEFAVINISSEPYFG 480
 DB 421 vwvgsvntvagmvekpfenisvkvlpkplnapnvldthgnefavinisseyfg 480

QY 473 -----EPKS-----CDKTHCPGPCPAPPELLGGP-----S 496
 DB 473 -----EPKS-----CDKTHCPGPCPAPPELLGGP-----S 496

QY 481 ILYKPVQAWKYIEVTEITLNYLEPRDYELC--VQLARPGEGEQHPQVRRITLAC 538
 DB 481 ilypvqawkyievteltnyleprdyelc--vqlarpggeqhpqpvrrittac 538

QY 497 VFLPFPKPKDTLMTLSRTP-EVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTPRREQYNS 555
 DB 497 vflpfpkpkdtlmtlsrtp-evtcvvvdvshedpevkfnwnyvdgvevhnaaktprreqyns 555

QY 539 IGLPPRGLSILPKSQALNTWQPIFNSED-----CFY----VEVCRSLQTTSDGQN- 589
 DB 539 iglpprglsilpksqalntwqpfinsed-----cfy----vevcrslqtsdgn- 589

QY 556 TYPVSVLIT-VLHQFWLWLNCKEYFKVSNKALPAPIERTISKAKGQ-PRPEQVVT----LP 609
 DB 556 typvsvlit-vlhwfwlwlnckeyfkvsnkallpapiertiskakgq-prpeqvvt----lp 609

QY 590 IKYVGNLTSVLSNLYPQGYTVRA-----RVNTKADGEWSCELRATWLSLILP 638
 DB 590 ikvgnltsvlsnlypqgytvra-----rvntkadgewseclratwlsdilp 638

QY 610 PSREEM-----PKNOVSLTCLVKCHYPSPDIAVWESNCOQPNY 648
 DB 610 psreem-----pknovsltclvkchypspdiahwesncqpny 648

QY 639 PGPENIKSLNITDSTANVSWL-IVDQYSISSIIIRYKVAQKNEDQH 683
 DB 639 ppenikslnitdstanvswl-ivdqysissiiirykvaqknedqh 683

QY 556 TVHVSQVIT VLRQWINGKFKYKAVSNKALAPLEKILSKAKGQ-PRKPVYI-----LP 609
 Db 590 -IKVpnllsvllsvlpvreqvltva-----rvntkaqewseelrwlldslp 638
 QY 610 PSREEM-----TKNQSILCLVKGYPYSLIAVWESNGQPPENNY 648
 Db 639 pipenikisnltstamswt-ivdgysissitrykvqgknedgh 683

RESULT 12

AAR53146
 ID AAR53146 standard; Protein; 1117 AA.

AC AAR53146;

XX 21-NOV-1994 (first entry)

XX Mouse tyrosine kinase ("tek") expressed during cardiogenesis.

XX Endothelial specific receptor tyrosine kinase; RTK; tek gene;

KW mouse; murine; chromosome 4; cardiogenesis

XX Mus musculus.

XX Key location/Qualifiers

FT Domain 821..1117

FT /label- catalytic_domain

FT Region 821..829

FT /label- conserved_region_I

FT /note- "the tek amino acid sequence appears elsewhere in the specification with the residues FQDV between positions 822 and 823"

FT Region 840..857

FT /label- conserved_region_II

FT Binding-site 823..829

FT /note- "GXGXC motif,

FT Binding-site 846..848 part of consensus ATP binding site"

FT /note- "AXK motif,

FT part of consensus ATP-binding site"

FT Region 858..873

FT /label- conserved_region_III

FT Region 874..888

FT /label- conserved_region_IV

FT Region 889..907

FT /label- conserved_region_V

FT Region 908..928

FT /label- insert

FT Region 929..969

FT /label- conserved_region_VI

FT Region 970..988

FT /label- conserved_region_VII

FT Region 975..977

FT /note- "motif found in almost all known kinases"

FT Region 989..1003

FT /label- conserved_region_VIII

FT Region 997..1003

FT /note- "motif found in transmembrane RTKs"

FT Region 1004..1034

FT /label- conserved_region_IX

FT Region 1035..1055

FT /label- conserved_region_X

FT /note- "the tek amino acid sequence appears elsewhere in the specification with the residues EL at positions 1039-1040 rather than DV"

FT Region 1056..1085

FT /label- conserved_region_XI

FT Region 1086..1117

FT /label- conserved_region_XII

XX

PN CA2085291-A.

XX 31-JAN-1994.

XX 14-DEC-1992; 92/A-2085291.

XX 30-JUL-1992; 92US-0921795.

XX (MOUN) MOUNT SINAI HOSPITAL CORP.

XX Breilman ML, Dumont DJ, Rossant J, Yamadauchi TP;

XX WPI; 1994-126938/16.

XX N-PSDB; AAQ62129.

XX Receptor tyrosine kinase isolated from cells of the endothelial

PT lineage - used for identifying (antagonists of ligand-receptor

PT binding

XX Claim 9; Fig 1; 5pp; English.

XX PNA from 9.5 and 12.5 day old mice was used to synthesize cDNA by RT

CC PCR using primers previously demonstrated to preferentially amplify

CC tyrosine kinase sequences. Clones correspond to 5 distinct tyrosine

CC kinases expressed during murine cardiogenesis were identified. Four

CC which were designated "tek". AAR53146 is the deduced tek amino acid

CC sequence. Comparisons to other tyrosine kinases showed that the

CC putative kinase domain contains several conserved sequence motifs

CC (see features table).

XX Sequence 1117 AA;

SQ

Query Match 62.4%; Score 2412; DB 15; Length 1117;

Best Local Similarity 67.8%; Pred. No. 3,60-137;

Matches 479; Conservative 50; Mismatches 95; Indels 82; Gaps 15;

QY 1 MDSLASLVAGVSLLSGIVEGMDLILINSILPLVSDAEISLTGASGRPHIEPTIGRD 60

Db 1 mdsiaqlvcvsllyqvveqamdillinsilplvsdaelsltcisaqwhphieptiqr 60

QY 61 FPAIMNHQDPIEVQWVIERWAKVVMKPKASKINCAVECEGVKPEFAIRPTMQRQ 120

Db 61 fealmnqhdpdvleqdrwakkvwwkaskiaqyiceqrqrqrirrtamqrq 120

QY 121 QASFLPATITMTVDKGDWNTSFRKVLKEEDAVIYKNGSIHSPRIEVPDILEVHLPH 180

Db 121 qasflpatitmtvdrdgvnvisfkkvlikee-davlykngsfihsvpihcvpdlievhlph 180

QY 181 AQPQDAGVYSARYIGGNLFTSAFTLIVKKEIAKWKPEINHLTACMNGVCHIEDTGE 240

Db 181 aqpdaqvysarylqgnlftsaftlrvrceadkwpdserpctcknaavchedtger 240

QY 241 ITPDPEMPTCEKAFITFGPTCKECSGCPCKSVVFCPLPWPYVCSAIVKNGKQCNF 300

Db 241 icppaimartekacephlfarlcke asapeqcksyvclpdpdyqscsalqzqlqne 299

QY 301 AHPFYGPQKLRKSPNNWEMTHFPWGLSPQWQHLQVPEPPIPMPTKIVDLPDIE 360

Db 300 acpsgyygpdccklrchctneeicdrfgqclscgqgqglqckeckgrpmtpdiedlpdhe 359

QY 361 VNSGKFNPTCKASWPLPTNFMTLVKPDGTVLHKPKNHTCHESVALFTTHPTLPDSG 420

Db 360 vnsqkfnptckaswplptseemlvkpdgvlrpdgltvdrpdltvdrpdltvdrpdlsq 419

QY 421 VWQSVNTVAGMVPKPPNI SVKVI PKPI NAI NVINIDTGHFAVINISSEPPYFG----- 472

Db 420 vwqsvntvagmvpknpsvkvylpeplhapavidchntalnisseyplqadpikskk 479

QY 473 -----EPKS-----CDKTHITPPCPAPPELLGAP-----S 496

Db 480 ifkpyndawkyievtnciftnlyieprtdyele-vglarpgqgqgqgqgqvrrttar 537

1. The first part of the document discusses the importance of maintaining accurate records of all transactions and the role of the accounting system in providing reliable financial information. It emphasizes the need for transparency and accountability in financial reporting.

2. The second part of the document outlines the various components of the accounting system, including the general ledger, subsidiary ledgers, and the trial balance. It explains how these components work together to ensure the accuracy and integrity of the financial data.

3. The third part of the document focuses on the process of reconciling bank statements with the company's records. It provides a step-by-step guide to identifying and resolving discrepancies, ensuring that the company's financial records are up-to-date and accurate.

4. The fourth part of the document discusses the importance of internal controls in preventing fraud and errors. It outlines the key elements of a strong internal control system, such as segregation of duties, authorization procedures, and regular audits.

5. The fifth part of the document addresses the challenges of managing financial data in a complex and rapidly changing business environment. It offers strategies for staying organized, efficient, and compliant with evolving regulations and standards.

6. The sixth part of the document provides a summary of the key points discussed throughout the document. It reiterates the importance of accurate record-keeping, transparency, and internal controls in ensuring the reliability of financial information.

7. The seventh part of the document includes a list of references and resources for further study. It points to various books, articles, and online resources that provide additional information on accounting principles and practices.

8. The eighth part of the document is a conclusion that summarizes the overall message of the document. It emphasizes the importance of maintaining accurate financial records and the role of the accounting system in providing reliable information for decision-making.

9. The ninth part of the document is a list of appendices that provide additional information and data. These appendices include sample forms, charts, and tables that illustrate the concepts discussed in the main text.

10. The tenth part of the document is a list of footnotes that provide additional information and references. These footnotes are used to provide more detail on specific points discussed in the main text and to cite the sources of the information used in the document.

GenCore version 4.5
Copyright (c) 1993 - 2000 Genengen Inc.

OM protein - protein search, using sw model

Run on: September 23, 2002, 22:03:19 : Search time 25.42 seconds
(without alignments)
676.460 Million cell updates/sec

Title: US-09-733-764-2
Perfect score: 3868
Sequence: 1 MSLASLVCGVSLLSGTV. : MIEALHNIYTKSLSLSPGK 764

Scoring table: HLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/1/aa/5A_COMB.pep.*
2: /cgn2_6/prodata/1/aa/5B_COMB.pep.*
3: /cgn2_6/prodata/1/aa/6A_COMB.pep.*
4: /cgn2_6/prodata/1/aa/6B_COMB.pep.*
5: /cgn2_6/prodata/1/aa/PC13S_COMB.pep.*
6: /cgn2_6/prodata/1/aa/bank11es1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2623	67.8	1124	1	US-08-323-474-2
2	2623	67.8	1124	2	US-09-296-005-15
3	2427.5	62.8	1118	1	US-07-934-393B-2
4	2427.5	62.8	1118	1	US-08-378-089A-2
5	2427.5	62.8	1118	2	US-08-838-957A-2
6	2427.5	62.8	1122	2	US-08-278-089A-6
7	2427.5	62.8	1122	2	US-08-838-957A-6
8	2488.5	61.8	1131	2	US-08-463-537A-6
9	1311	33.9	911	2	US-08-484-438-10
10	1300	33.6	918	1	US-08-385-229-4
11	1298.5	33.6	486	1	US-08-243-010-1
12	1279.5	33.1	396	2	US-08-784-512-3
13	1279.5	33.1	396	2	US-09-176-228-3
14	1276.5	33.0	438	1	US-08-047-827-11
15	1276.5	33.0	438	1	US-08-404-574-11
16	1274	32.9	424	5	US-09-095-03866-12
17	1274	32.9	424	5	US-09-095-03866-14
18	1268	32.8	451	2	US-08-887-352B-14
19	1268	32.8	451	2	US-08-287-452B-14
20	1268	32.8	451	2	US-08-887-352B-18
21	1268	32.8	451	3	US-08-466-151-65
22	1268	32.8	451	4	US-09-109-207C-14
23	1268	32.8	451	4	US-09-109-207C-16
24	1268	32.8	451	4	US-09-109-207C-18
25	1268	32.8	451	4	US-09-282-505-2
26	1268	32.8	451	4	US-09-864-206-2
27	1268	32.8	451	4	US-09-296-005-14

Sequence 16, Appl
Sequence 18, Appl
Sequence 71, Appl
Sequence 71, Appl
Sequence 8, Appl
Sequence 8, Appl
Sequence 22, Appl
Sequence 22, Appl
Sequence 22, Appl
Sequence 4, Appl
Sequence 18, Appl
Sequence 18, Appl
Sequence 9, Appl
Sequence 11, Appl
Sequence 2, Appl
Sequence 2, Appl
Sequence 15, Appl

ALIGNMENTS

RESULT 1
US-08-323-474-2
: Sequence 2, Application US/08323474
: Patent No. 5447860
: GENERAL INFORMATION:
: APPLICANT: Ziegler, Steven F.
: TITLE OF INVENTION: NOVEL TYROSINE KINASE
: NUMBER OF SEQUENCES: 8
: CORRESPONDENCE ADDRESS:
: ADDRESS: Immunex Corporation
: STREET: 51 University Street
: CITY: Seattle
: STATE: Washington
: COUNTRY: US
: ZIP: 98101
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/323.474
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/905,600
: FILING DATE: 26 JUN 1992
: ATTORNEY/AGENT INFORMATION:
: NAME: Seese, Kathryn A.
: REGISTRATION NUMBER: 32,172
: REFERENCE/EXHIBIT NUMBER: 2509
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (206) 587-0430
: TELEFAX: (206) 233-0644
: TELEX: 756822
: INFORMATION FOR SEQ ID NO. 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1124 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MODIFICATION TYPE: protein

US 08-323-474-2

Query Match: 67.8%, Score 2623, Lb 1, Length 1124;
Best local Similarity: 77.18; Pred No. 2, 80-199;
Matches 501, Conservative 23, Mismatches 44; Indels 82; Gaps 7;
QY : MSLASLVCGVSLLSGTVSAMLLIINSFLVSLALISLCLASQWRHEITIGED 60
|||||

```

Db 1 MDLALSLVLSVSLLSSTVEGAMILLINSLPLVSDAETSLTCTASGWRPHIEITIGRD 60
QY 61 FPAIMNQHODPLEVTDVIRWAKKVVWKKPKASK INCAVEFCGVRGFAIPITMKMPQ 120
Db 63 FPAIMNQHODPLEVTDVIRWAKKVVWKKPKASK INCAVEFCGVRGFAIPITMKMPQ 120
QY 121 QASFLPATLTMTVDKGDVNIISFKKVIKEEDAVIYKNGSFIHSVPRIEVPDILEVILPH 180
Db 121 QASFLPATLTMTVDKGDVNIISFKKVIKEEDAVIYKNGSFIHSVPRIEVPDILEVILPH 180
QY 181 AQOQACVVSARYIGCNLTSAFTPLIVRPEACQKWPENHILATACMNNVCHIEDTGC 240
Db 181 AQOQACVVSARYIGCNLTSAFTPLIVRPEACQKWPENHILATACMNNVCHIEDTGC 240
QY 241 ICPPGFMGRTCKACELHTPGRTCKERCSCGGCKSVYVFCPLDPYGCSCATGKGLQNE 300
Db 241 ICPPGFMGRTCKACELHTPGRTCKERCSCGGCKSVYVFCPLDPYGCSCATGKGLQNE 300
QY 301 ACHPEYFGPKKLRSCNNGEMCDRFQGLSPGWGLQTEREGIPRMTKIVDLPHIE 360
Db 301 ACHPEYFGPKKLRSCNNGEMCDRFQGLSPGWGLQTEREGIPRMTKIVDLPHIE 360
QY 361 VNSGFENPICKASWPLTNEEMTLVKPDGTVLHPKDNHTDHSVAIFTHRLPPDSG 420
Db 361 VNSGFENPICKASWPLTNEEMTLVKPDGTVLHPKDNHTDHSVAIFTHRLPPDSG 420
QY 421 VWCSTNTVAGVKEKPFNISVKVLKPLNAPNVIDTCHNFAVINISSEPYGEPKSCDKT 480
Db 421 VWCSTNTVAGVKEKPFNISVKVLKPLNAPNVIDTCHNFAVINISSEPYGEPKSCDKT 480
QY 481 HTCPCPAPPELLGSGSVFLFPKPKDITMISRTPEVTCVVVDVSHEDPEVKFNHYV 536
Db 474 ---GPIKSKKIL-----YKPVNHYEAMQHIOVINEIVTLNLEPRTEYELCVOLVR 521
QY 537 ---DRYVEVINAKTKPREQVNSTYRVVSVLTVLHQWLNGKEYCKVSNKALPAPIEKTIS 594
Db 539 ASIG-----LPPRGLNLLPKSQTLLNLTWQIFPSSDEDFYVEVE 579

RESULT 3
PCT-US93-06093-2
: Sequence 2, Application PC/IUS9306093
: GENERAL INFORMATION:
: APPLICANT: Ziegler, Steven F.
: TITLE OF INVENTION: NOVEL TYROSINE KINASE
: NUMBER OF SEQUENCES: 3
: CORRESPONDENCE ADDRESS:
: ADDRESS: Immunex Corporation
: STREET: 51 University Street
: CITY: Seattle
: STATE: Washington
: COUNTRY: US
: ZIP: 98101
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US93/06093
: FILING DATE: 19930625
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 01/905,600
: FILING DATE: 26-JUN-1992
: ATTORNEY/AGENCY INFORMATION:
: NAME: Sooco, Kathryn A.
: REGISTRATION NUMBER: 32,172
: REFERENCE/DOCKET NUMBER: 2609

```

```

: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (206) 587-0430
: TELEFAX: (206) 233-9644
: TELEX: 756822
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1124 amino acids
: TYPE: AMINO ACID
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: PCT-US93-06093-2

Query Match 67.8% Score 2623; DB 5; Length 1124;
Best Local Similarity 77.1%; Pred. No. 2,80-199;
Matches 501; Conservative 23; Mismatches 44; Indels 82; Gaps 7;

QY 1 MDLALSLVLSVSLLSSTVEGAMILLINSLPLVSDAETSLTCTASGWRPHIEITIGRD 60
Db 1 MDLALSLVLSVSLLSSTVEGAMILLINSLPLVSDAETSLTCTASGWRPHIEITIGRD 60
QY 61 FPAIMNQHODPLEVTDVIRWAKKVVWKKPKASK INCAVEFCGVRGFAIPITMKMPQ 120
Db 61 FPAIMNQHODPLEVTDVIRWAKKVVWKKPKASK INCAVEFCGVRGFAIPITMKMPQ 120
QY 121 QASFLPATLTMTVDKGDVNIISFKKVIKEEDAVIYKNGSFIHSVPRIEVPDILEVILPH 180
Db 121 QASFLPATLTMTVDKGDVNIISFKKVIKEEDAVIYKNGSFIHSVPRIEVPDILEVILPH 180
QY 181 AQOQACVVSARYIGCNLTSAFTPLIVRPEACQKWPENHILATACMNNVCHIEDTGC 240
Db 181 AQOQACVVSARYIGCNLTSAFTPLIVRPEACQKWPENHILATACMNNVCHIEDTGC 240
QY 241 ICPPGFMGRTCKACELHTPGRTCKERCSCGGCKSVYVFCPLDPYGCSCATGKGLQNE 300
Db 241 ICPPGFMGRTCKACELHTPGRTCKERCSCGGCKSVYVFCPLDPYGCSCATGKGLQNE 300
QY 301 ACHPEYFGPKKLRSCNNGEMCDRFQGLSPGWGLQTEREGIPRMTKIVDLPHIE 360
Db 301 ACHPEYFGPKKLRSCNNGEMCDRFQGLSPGWGLQTEREGIPRMTKIVDLPHIE 360
QY 361 VNSGFENPICKASWPLTNEEMTLVKPDGTVLHPKDNHTDHSVAIFTHRLPPDSG 420
Db 361 VNSGFENPICKASWPLTNEEMTLVKPDGTVLHPKDNHTDHSVAIFTHRLPPDSG 420
QY 421 VWCSTNTVAGVKEKPFNISVKVLKPLNAPNVIDTCHNFAVINISSEPYGEPKSCDKT 480
Db 421 VWCSTNTVAGVKEKPFNISVKVLKPLNAPNVIDTCHNFAVINISSEPYGEPKSCDKT 480
QY 481 HTCPCPAPPELLGSGSVFLFPKPKDITMISRTPEVTCVVVDVSHEDPEVKFNHYV 536
Db 474 ---GPIKSKKIL-----YKPVNHYEAMQHIOVINEIVTLNLEPRTEYELCVOLVR 521
QY 537 ---DRYVEVINAKTKPREQVNSTYRVVSVLTVLHQWLNGKEYCKVSNKALPAPIEKTIS 594
Db 539 ASIG-----LPPRGLNLLPKSQTLLNLTWQIFPSSDEDFYVEVE 579

RESULT 3
: Sequence 2, Application US/07934393H
: Patent No. 5466596
: GENERAL INFORMATION:
: APPLICANT: BREITMAN, MARTIN L.
: APPLICANT: DUMONT, DANIEL
: APPLICANT: GRAUHOHL, GERARD G.
: TITLE OF INVENTION: TISSUE SPECIFIC TRANSCRIPTIONAL
: REGULATORY ELEMENT
: NUMBER OF SEQUENCES: 5

```

; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: BERESKIN & PARR
 ; STREET: 40 King Street West
 ; CITY: Toronto
 ; STATE: Ontario
 ; COUNTRY: Canada
 ; ZIP: M5H 3Y2
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/934,393R
 ; FILING DATE: 25-AUG-1992
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Kurdydyk, Linda M.
 ; REGISTRATION NUMBER: 34,971
 ; REFERENCE/DOCKET NUMBER: 3153-64
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (416) 354-7311
 ; TELEFAX: (416) 361-1398
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1118 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-07-934-393B-2

Query Match 62.8%; Score 2427.5; DB 1; Length 1118;
 Best Local Similarity 68.0%; Pred. No. 8.3e-184;
 Matches 480; Conservative 50; Mismatches 95; Indels 81; Gaps 14,
 1 MDSLASLVLCGVSLLSGTVEGAMDLILNSLPVSDAETSLTCTASGPPHEPTIGRD 60
 1 MDSLAGLVLCGVSLLSGTVEGAMDLILNSLPVSDAETSLTCTASGPPHEPTIGRD 60
 61 FEALMNOHQDPLEVTQDVTPEWAKKVVWPEKASKINGAYFCGPGVGRGAIPRIMKMPQ 120
 61 FEALMNOHQDPLEVTQDVTPEWAKKVVWPEKASKINGAYFCGPGVGRGAIPRIMKMPQ 120
 121 QASFLPATLITMTVDKGDVNIISFKKVLKEEDAVIYKNGSFHSVPHEVDPDILEVHLPH 180
 121 QASFLPATLITMTVDKGDVNIISFKKVLKEEDAVIYKNGSFHSVPHEVDPDILEVHLPH 180
 181 AOPQDAVYSARYTGSNLTSAETPLIVRCEAQKWPCECNHLCCTAMNNGVCHEDTIGEC 240
 181 AOPQDAVYSARYTGSNLTSAETPLIVRCEAQKWPCECNHLCCTAMNNGVCHEDTIGEC 240
 241 ICPGFGMRKTCCKATPHITGPTCKPMSGDPKSKSVVPCIPHPYGSATAGSKGLQCN 300
 241 ICPGFGMRKTCCKATPHITGPTCKPMSGDPKSKSVVPCIPHPYGSATAGSKGLQCN 300
 301 ACPHCPGPPCKICPCSNNGHMCDFGCGICSPGSGGLQCEPFCIPPHITKIVULPHIE 360
 301 ACPHCPGPPCKICPCSNNGHMCDFGCGICSPGSGGLQCEPFCIPPHITKIVULPHIE 360
 361 VNSGKFNPKCKASGMPLEPTNEMLTVKPDGTVLHPKDFNHTDHFSAIFTHIRLPPDSG 420
 361 VNSGKFNPKCKASGMPLEPTNEMLTVKPDGTVLHPKDFNHTDHFSAIFTHIRLPPDSG 420
 421 VVVCSSNTVAGMVEKPNISVKVLKPLNAPNVIDTCHNFAVINISSEPYFG----- 472
 421 VVVCSSNTVAGMVEKPNISVKVLKPLNAPNVIDTCHNFAVINISSEPYFG----- 472
 473 -----EPKS-----CDKTHTCPPCAPRLGLCP-----S 496
 481 LFYKPVNQAWKYIEVTNFIPLTNLEPRTDYELC--VOLARPGGGCGCHCPVRRFTTAC 538
 497 VFLPPPKDGLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVGVVFIHNAIKTPREEQYNS 555

Db 539 IGLPPKGLSLLPKSQAIALNLTWQPIFTNSD-----EPY---VEVERSLQITISDQON- 589
 QY 556 TYRVVSVLT-VLRDWMNGCKYKCKVSNKALPAPIKHTISKAKGO-PREPOVYI---LP 609
 Db 590 -IKVFNLTSLVLSNLVPEQYTVFA-----PNTKAGGSESELRAWLTLSLILP 638
 QY 610 PSPEEM-----TRNQVSLTCLVKGFYPSDIAVWESNGQPENNY 648
 Db 639 PPHENIKISNIIHSTAMVSWT-IVVQYSISSIIHKYQCKKNFVGH 683

RESULT 4

US-08-278-089A-2
 ; Sequence 2, Application US/98278089A
 ; Patent No. 5681714
 ; GENERAL INFORMATION:
 ; APPLICANT: Breilman, Martin L.
 ; APPLICANT: Rossant, Janet
 ; APPLICANT: Dumont, Daniel J.
 ; APPLICANT: Yamaguchi, Terry P.
 ; TITLE OF INVENTION: NO. 5681714e1 Receptor Tyrosine Kinase
 ; NUMBER OF SEQUENCES: 33
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Bereskin & Parr
 ; STREET: 40 King Street West
 ; CITY: Toronto
 ; STATE: Ontario
 ; COUNTRY: Canada
 ; ZIP: M5H 3Y2
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/98/278,089A
 ; FILING DATE: 20-JUL-1994
 ; CLASSIFICATION: 530
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Kurdydyk, Linda M.
 ; REGISTRATION NUMBER: 34,971
 ; REFERENCE/DOCKET NUMBER: 3153-111
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (416) 364-7311
 ; TELEFAX: (416) 361-1398
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1118 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-278-089A-2

Query Match 62.8%; Score 2427.5; DB 1; Length 1118;
 Best Local Similarity 68.0%; Pred. No. 8.3e-184;
 Matches 480; Conservative 50; Mismatches 95; Indels 81; Gaps 14,
 1 MDSLASLVLCGVSLLSGTVEGAMDLILNSLPVSDAETSLTCTASGPPHEPTIGRD 60
 1 MDSLAGLVLCGVSLLSGTVEGAMDLILNSLPVSDAETSLTCTASGPPHEPTIGRD 60
 61 FEALMNOHQDPLEVTQDVTPEWAKKVVWPEKASKINGAYFCGPGVGRGAIPRIMKMPQ 120
 61 FEALMNOHQDPLEVTQDVTPEWAKKVVWPEKASKINGAYFCGPGVGRGAIPRIMKMPQ 120
 121 QASFLPATLITMTVDKGDVNIISFKKVLKEEDAVIYKNGSFHSVPHEVDPDILEVHLPH 180
 121 QASFLPATLITMTVDKGDVNIISFKKVLKEEDAVIYKNGSFHSVPHEVDPDILEVHLPH 180
 181 AOPQDAVYSARYTGSNLTSAETPLIVRCEAQKWPCECNHLCCTAMNNGVCHEDTIGEC 240
 181 AOPQDAVYSARYTGSNLTSAETPLIVRCEAQKWPCECNHLCCTAMNNGVCHEDTIGEC 240

```

Db 181 AQPQAGVYSARYIGGNLTSAFTLIVRRGIAQKQWGHGSRPCTCKNNQVCHEDTGEC 240
QY 241 IGPPPEPMGRGCEKATETITETGTCTEPASQEDGSKSVWFLPDPYGSASATGNKGLAWNE 300
Db 241 ITPPPPMGRGTCERAKTEPHITFTPTPEPSGEPFGSKSVWFLPDPYGSASATGNKGLAWNE 300
QY 301 ACHPGYGPDPCKLRGSCNNGMCDRFQGCGLGSPGWOGIQCEREGIPRMTPKIVLDPDPIIE 360
Db 301 ACPSPGYGPDCKLPQWTFNFTCTDPFGGLASQAWQGLQCEKEGPPPMTPQIEDLDPDPIIE 360
QY 361 VNSGFNPICAKSGWPLTNFMTLVKPDGTVLHPKDFNIHDSVAIFTHIRILPPDSG 420
Db 361 VNSGFNPICAKSGWPLTNFMTLVKPDGTVLHPKDFNIHDSVAIFTHIRILPPDSG 420
QY 421 VVCSVNTVAGVWEKPFENISVKVLPKPLNAPNVDITGHNFAVINISSEPFVFG----- 472
Db 421 VVCSVNTVAGVWEKPFENISVKVLPKPLNAPNVDITGHNFAVINISSEPFVFGDGPIKSKK 480
QY 473 -----EPKS-----CQKTHCTCPCHAPHELGGP-----S 496
Db 481 LEYKPVNQAWKYIEVTNEIFTLNLEPKTDYELC--VOLAKPGEGGHHGPIVRRRTTAC 548
QY 497 VFLFPKPKDGLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNS 555
Db 539 IGLPPHPRGLSLPKSOTALNLTWQPIFTNSHD-----EFY---VEVEKRKRSIQTTSDQON- 589
QY 556 TYRVVSVLT-VLIHQDLNCKEYKCKVSNKALPAPIETKISKAKGQ-PRPQVYT-----LP 609
Db 590 -IKVPGNLTSLVLSNLPVREQTVRA-----RVNTKAUGSESEELRAWTLSILP 638
QY 610 PSREEM-----TKNQVSLTCLVKGFYPSDIAVEHWSNCOGPNENY 648
Db 639 PGPENIKTSNITUSTAMVSWT-IVNGYSISSIIIRYKVGKQKNEQDH 683

```

RESULT 5

```

US-08-838-957A-2
Sequence 2, Application US/08/838/957A
Patent No. 5998187
GENERAL INFORMATION:
APPLICANT: Breitman, Martin L.
APPLICANT: Rossant, Janet
APPLICANT: Dumont, Daniel J.
APPLICANT: Yamaguchi, Terry P.
TITLE OF INVENTION: No. 5998187el Receptor Tyrosine Kinase
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSER: Bereskin & Parr
STREET: 40 King Street West
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5H 3Y2
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/838/957A
FILING DATE: 23-APR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kurdydyk, Linda M.
REGISTRATION NUMBER: 34,971
REFERENCE/DOCKET NUMBER: 3153-212
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 364-7311
TELEFAX: (416) 361-1398
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1118 amino acids
TYPE: amino acid

```

```

; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-838-957A-2
Query Match 62.8%; Score 2427.5; DB 2; Length 1118;
Best Local Similarity 68.0%; Pred. No. 8,30-184;
Matches 480; Conservative 50; Mismatches 95; Indels 81; Gaps 14;
QY 1 MSLASLVLCGVSLLSGLVEGAMULLLINSILPLYSIAETLSLTGASGWRPHIEPTITGRD 60
Db 1 MSLASLVLCGVSLLSGLVEGAMULLLINSILPLYSIAETLSLTGASGWRPHIEPTITGRD 60
QY 61 FEALNQHQDDPLEVTQDVITREWAKKVVWKKKASKINCAVFCGKRVGCEALRIKIMKMO 120
Db 61 FEALNQHQDDPLEVTQDVITREWAKKVVWKKKASKINCAVFCGKRVGCEALRIKIMKMO 120
QY 121 QASFLPATLITMTVDKGNVNISPKKVLIKEDAVIYKNGSPFHSVPKRVHVDLLEVLPH 180
Db 121 QASFLPATLITMTVDKGNVNISPKKVLIKEDAVIYKNGSPFHSVPKRVHVDLLEVLPH 180
QY 181 AQPQAGVYSARYIGGNLTSAFTLIVRRGIAQKQWGHGSRPCTCKNNQVCHEDTGEC 240
Db 181 AQPQAGVYSARYIGGNLTSAFTLIVRRGIAQKQWGHGSRPCTCKNNQVCHEDTGEC 240
QY 241 IGPPPEPMGRGCEKATETITETGTCTEPASQEDGSKSVWFLPDPYGSASATGNKGLAWNE 300
Db 241 ITPPPPMGRGTCERAKTEPHITFTPTPEPSGEPFGSKSVWFLPDPYGSASATGNKGLAWNE 300
QY 301 ACHPGYGPDPCKLRGSCNNGMCDRFQGCGLGSPGWOGIQCEREGIPRMTPKIVLDPDPIIE 360
Db 301 ACPSPGYGPDCKLPQWTFNFTCTDPFGGLASQAWQGLQCEKEGPPPMTPQIEDLDPDPIIE 360
QY 361 VNSGFNPICAKSGWPLTNFMTLVKPDGTVLHPKDFNIHDSVAIFTHIRILPPDSG 420
Db 361 VNSGFNPICAKSGWPLTNFMTLVKPDGTVLHPKDFNIHDSVAIFTHIRILPPDSG 420
QY 421 VVCSVNTVAGVWEKPFENISVKVLPKPLNAPNVDITGHNFAVINISSEPFVFG----- 472
Db 421 VVCSVNTVAGVWEKPFENISVKVLPKPLNAPNVDITGHNFAVINISSEPFVFGDGPIKSKK 480
QY 473 -----EPKS-----CQKTHCTCPCHAPHELGGP-----S 496
Db 481 LEYKPVNQAWKYIEVTNEIFTLNLEPKTDYELC--VOLAKPGEGGHHGPIVRRRTTAC 548
QY 497 VFLFPKPKDGLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNS 555
Db 539 IGLPPHPRGLSLPKSOTALNLTWQPIFTNSHD-----EFY---VEVEKRKRSIQTTSDQON- 589
QY 556 TYRVVSVLT-VLIHQDLNCKEYKCKVSNKALPAPIETKISKAKGQ-PRPQVYT-----LP 609
Db 590 -IKVPGNLTSLVLSNLPVREQTVRA-----RVNTKAUGSESEELRAWTLSILP 648
QY 610 PSREEM-----TKNQVSLTCLVKGFYPSDIAVEHWSNCOGPNENY 648
Db 639 PGPENIKTSNITUSTAMVSWT-IVNGYSISSIIIRYKVGKQKNEQDH 683

```

RESULT 6

```

US-08-278-089A-6
Sequence 6, Application US/08278089A
Patent No. 5681714
GENERAL INFORMATION:
APPLICANT: Breitman, Martin L.
APPLICANT: Rossant, Janet
APPLICANT: Dumont, Daniel J.
APPLICANT: Yamaguchi, Terry P.
TITLE OF INVENTION: No. 5681714el Receptor Tyrosine Kinase
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSER: Bereskin & Parr
STREET: 40 King Street West
CITY: Toronto

```

STATE: Ontario
 COUNTRY: Canada
 ZIP: M5H 3Y2
 COMPUTER PEATABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/278,089A
 FILING DATE: 20-Jul-1994
 CLASSIFICATION: 530
 ATTORNEY/AGENT INFORMATION:
 NAME: Kurdydyk, Linda M.
 REGISTRATION NUMBER: 34,971
 REFERENCE/DOCKET NUMBER: 3153-111
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (416) 361-1398
 TELEFAX: (416) 361-1398
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1122 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-278-089A-6

Query Match 62.8%; Score 2427.5; DB 1; Length 1122;
 Best Local Similarity 58.0%; Pred. No. 8.4e-184;
 Matches 480; Conservative 50; Mismatches 95; Indels 81; Gaps 14;

QY 1 MDSLASLVLCGVSLLESSTVERAMGILLINSLPLVSDAETSLTCLASGWRPHEPTIGRD 60
 DB 1 MDSLAGLVLCGVSLLYGVYVGGAMDELLINSLPLVSDAETSLTCLASGWRPHEPTIGRD 60
 QY 61 FEALMNHQHQPLEVTQDVTFRWAKKVVWKKKASKINGAYFCGVRGCAIRIRTKMKRQ 120
 DB 61 FEALMNHQHQPLEVTQDVTFRWAKKVVWKKKASKINGAYFCGVRGCAIRIRTKMKRQ 120
 QY 121 QASFLPATLMTVVGKGNVNISFKKVLKEEDAVIYKNGSFIRSVPRHEVFDILEVHLPH 180
 DB 121 QASFLPATLMTVVGKGNVNISFKKVLKEEDAVIYKNGSFIRSVPRHEVFDILEVHLPH 180
 QY 181 AQODAGVYSARYTGGNLTSAFTPLIVPCEAGKMGPECNHLCIACMNGVCHEDTGC 240
 DB 181 AQODAGVYSARYTGGNLTSAFTPLIVPCEAGKMGPECNHLCIACMNGVCHEDTGC 240
 QY 241 ICPPGFMGRTCEKACELHTFORICKECGCGCKSVFCLPDGPGYGCSCAIGWKGLQCNE 300
 DB 241 ICPPGFMGRTCEKACELHTFORICKECGCGCKSVFCLPDGPGYGCSCAIGWKGLQCNE 300
 QY 301 ACHPGFYGPCKLPSTNNKEMTDPPGSLCLSPWQHLQEPESIPPMTPKIVLFDHIE 360
 DB 301 ACPGSGYGPCKLKHCHCTNPEICDRFGCLCSQWQGLQCEKCEGRPMTPQIEDLPDHI 360
 QY 361 VNSGKFNPKICKASGWLPTNEETLVKPDGTVLHPKDFNITDHFSAVFIIRHLPPDSG 420
 DB 361 VNSGKFNPKICKASGWLPTNEETLVKPDGTVLHPKDFNITDHFSAVFIIRHLPPDSG 420
 QY 421 VWGCVNIVAGMVKPKPNISVKVLKPKLINAPNVITDGHNPVAVINISSEPYFG----- 472
 DB 421 VWGCVNIVAGMVKPKPNISVKVLKPKLINAPNVITDGHNPVAVINISSEPYFG----- 472
 QY 473 -----EPKS-----CDKTHCTCPGAPPELLGGP-----S 496
 DB 481 LFYKPVQAMKYIEVINEITLANTPEPTDYEIC--VOLARPGGGCHGCPVPRRTTAC 538
 QY 497 VFLEPPKPKDTLMSRTP-EVTCVVWDVSHEDPEVKFNWVDGVEVINAKTKPKEOYNS 555
 DB 539 IGLPPKGLSLPKSQIALNLITWQPIETNSD-----EFY---VEVEKRSLOTTSDQUN- 589
 QY 556 TYRVSVLIIVLHQLWNGREYKCKVSNKALPALEKILSKAKGQ+PREPVVYI---LTP 609

DB 590 -IKVPCNITSVLISNIVPQYIVVA-----PVNIKAGDSSEFLPAWILSLIP 638
 QY 610 PSREEM-----TKNOVSLTCLVKGFYPSDIAVWESNGOPPENNY 648
 DB 639 PQPNTKISNITDSTIAMVSWT-IVVGYSSISIIYKVGGRNETHQH 683

RESULT 7
 US-08-838-957A-6
 Sequence 5; Application US/08838957A
 Patent No. 5998187
 GENERAL INFORMATION:
 APPLICANT: Breitman, Martin L.
 APPLICANT: Kossant, Janet
 APPLICANT: Dumont, Daniel J.
 APPLICANT: Yamauchi, Terry P.
 TITLE OF INVENTION: No 5998187el Receptor Tyrosine Kinase
 NUMBER OF SEQUENCES: 32
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Bereskin & Parr
 STREET: 40 King Street West
 CITY: Toronto
 STATE: Ontario
 COUNTRY: Canada
 ZIP: M5H 3Y2
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #10, Version #1 30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/838-957A
 FILING DATE: 23-APR-1997
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Kurdydyk, Linda M.
 REGISTRATION NUMBER: 34,971
 REFERENCE/PATENT NUMBER: 3153-212
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (416) 364-7311
 TELEFAX: (416) 361-1398
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1122 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-838-957A-6

Query Match 62.8%; Score 2427.5; DB 2; Length 1122;
 Best Local Similarity 58.0%; Pred. No. 8.4e-184;
 Matches 480; Conservative 50; Mismatches 95; Indels 81; Gaps 14;

QY 1 MDSLASLVLCGVSLLSGTGVRGAMDELLINSLPLVSDAETSLTCLASGWRPHEPTIGRD 60
 DB 1 MDSLASLVLCGVSLLYGVYVGGAMDELLINSLPLVSDAETSLTCLASGWRPHEPTIGRD 60
 QY 61 FEALMNHQHQPLEVTQDVTFRWAKKVVWKKKASKINGAYFCGVRGCAIRIRTKMKRQ 120
 DB 61 FEALMNHQHQPLEVTQDVTFRWAKKVVWKKKASKINGAYFCGVRGCAIRIRTKMKRQ 120
 QY 121 QASFLPATLMTVVGKGNVNISFKKVLKEEDAVIYKNGSFIRSVPRHEVFDILEVHLPH 180
 DB 121 QASFLPATLMTVVGKGNVNISFKKVLKEEDAVIYKNGSFIRSVPRHEVFDILEVHLPH 180
 QY 181 AQODAGVYSARYTGGNLTSAFTPLIVPCEAGKMGPECNHLCIACMNGVCHEDTGC 240
 DB 181 AQODAGVYSARYTGGNLTSAFTPLIVPCEAGKMGPECNHLCIACMNGVCHEDTGC 240
 QY 241 ICPPGFMGRTCEKACELHTFORICKECGCGCKSVFCLPDGPGYGCSCAIGWKGLQCNE 300
 DB 241 ICPPGFMGRTCEKACELHTFORICKECGCGCKSVFCLPDGPGYGCSCAIGWKGLQCNE 300

```

Db 241 ICPGPMCKTCKACGPHFGRICKERKSGEGEGCKSVYVFLDPPYQSGSCARWHRGLQCNF 300
QY 301 ACBPGEFYRPUCKI RQSONNK:EMWDPFGCTASPGWGLQV:EPGIGPRMTPKIVLDLPDHE 360
Db 301 ACBPGEFYRPUCKI RQSONNK:EMWDPFGCTASPGWGLQV:EPGIGPRMTPKIVLDLPDHE 360
QY 361 VNSGRFNPICKASGWPLOINRMILVAKVGGIVLHFKDQFNHIDHFSVAIFTHKILPPDSG 420
Db 361 VNSGRFNPICKASGWPLOINRMILVAKVGGIVLHFKDQFNHIDHFSVAIFTHKILPPDSG 420
QY 421 VVWCVNTVAGWVKEPFIKSVKVLKPLNADPNDVTDGHNFAVINISSEPEFG----- 472
Db 421 VVWCVNTVAGWVKEPFIKSVKVLKPLNADPNDVTDGHNFAVINISSEPEFG----- 472
QY 473 -----EPKS CDKTHICPPCPAPELLGCP-----S 496
Db 481 LFYKPVNQAKWYIEVTNEIFTNILEPTDYELG--VOLAPP-EGEGEHP-PPVFFTHA 538
QY 497 VFLPPIKPKDILMISKIP-EVICVVVDVSHEDPEVK:NNVYDGVVEVHNAKIKPKEEQYNS 555
Db 539 IGLPPRGLSLIPKSGTALNLTWQPIETNSD-----EFY---VEVEPRSLQTTSDQCN- 589
QY 556 TYVWVSALT-VLHEDWLN:KFKVKSNAKLPAPIEVTSKAKQ:PPPEQVYT-----IP 609
Db 590 -IKVGNITSVIISNIVPEQVTHA-----PVNTKAG:FWSEETPAWTLSDLP 638
QY 610 PSREEM-----TKNQSLSLCKVGEPSDIAVENESNCPENNY 648
Db 639 PQPENIKISNITSDAMVSWT IVLGYSSTSIIRYKVKQKNEQH 664

```

RESULT R

US-08-469-537A-96

: Sequence 96, Application US/08469537A

: Patent No. 5843749

: GENERAL INFORMATION:

: APPLICANT: Maisonnier, et al.

: TITLE OF INVENTION: HKK AND FOR TYROSINE

: TITLE OF INVENTION: KINASES

: NUMBER OF SEQUENCES: 107

: CORRESPONDENCE ADDRESS:

: ADDRESSER: Regeneron Pharmaceuticals, Inc.

: STREET: 777 Old Saw Mill River Road

: CITY: Tarrytown

: STATE: NY

: COUNTRY: U.S.A.

: ZIP: 10591

: COMPUTER READABLE FORM:

: MEDIUM TYPE: Diskette

: COMPUTER: IBM Compatible

: OPERATING SYSTEM: DOS

: SOFTWARE: FastSeq Version 2.0

: CURRENT APPLICATION DATA:

: APPLICATION NUMBER: US/08/469,537A

: FILING DATE: 06-JUN-1995

: CLASSIFICATION: 435

: PRIOR APPLICATION DATA:

: APPLICATION NUMBER: USSN 08/406,247

: FILING DATE: 17-MAR-1995

: APPLICATION NUMBER: USSN 08/144,992

: FILING DATE: 28-OCT-1993

: APPLICATION NUMBER: USSN 07/736,559

: FILING DATE: 26-JUL-1991

: ATTORNEY/AGENT INFORMATION:

: NAME: Kopleff, Ph.D., Gail M

: REGISTRATION NUMBER: 32,143

: REFERENCE/DOCKET NUMBER: RFG 070C

: TELECOMMUNICATION INFORMATION:

: TELEPHONE: 914-345-7400

: TELEFAX: 914-345-7721

: TELEX:

: INFORMATION FOR SEQ ID NO: 96:

: SEQUENCE CHARACTERISTICS:

```

: LENGTH: 1101 amino acids
: TYPE: amino acid
: STRANDEDNESS: unknown
: TOPOLOGY: unknown
: MOLECULE TYPE: protein
US-08-469-537A-96

```

```

Query Match 61.89; Score 2388.5; DB 2; Length 1101;
Best Local Similarity 69.28; Pred. No. 1e-180;
Matches 473; Conservative 40; Mismatches 96; Indels 75; Gaps 14;

```

```

QY 20 VEGAMDLILINSLPLVSDAETSLICASQWPPHEPITIGRDFEALMKQRPHEVLTQDVT 79
Db 1 VEGAMDLILINSLPLVSDAETSLICASQWPPHEPITIGRDFEALMKQRPHEVLTQDVT 60
QY 80 PFWAKVVKWPEKASKINAYFETEPVPECEATPTPTMPPEQASTPATITMTVDKGDNV 149
Db 61 PFWAKVVKWPEKASKINAYFETEPVPECEATPTPTMPPEQASTPATITMTVDKGDNV 120
QY 140 NLSFKVLIRKEDAVYKNGSEIHISVPEHEVPDILEVILPBAQPDAGVYSARYIGNLF 199
Db 121 NLSFKVLIRKEDAVYKNGSEIHISVPEHEVPDILEVILPBAQPDAGVYSARYIGNLF 180
QY 200 TSAFTPIVPRPFAQKWPDPGPHICTACMNNVCVCHEDIGSCIDPPPCMPTCKAKELPT 259
Db 181 TSAFTPIVPRPFAQKWPDPGPHICTACMNNVCVCHEDIGSCIDPPPCMPTCKAKELPT 240
QY 260 PAPT-KKFG-SQPEKSKSVVECTIPQYKASCAIHWKGLQWENHPGEGYGPCKLRSCNN 419
Db 241 PAPT-KKFG-SQPEKSKSVVECTIPQYKASCAIHWKGLQWENHPGEGYGPCKLRSCNN 400
QY 320 SEMTDFQNTSPQWGLQWTEPESELPMPTPKIVLEPHIEVNSGKPNPCKASGWPPLPT 479
Db 301 SEMTDFQNTSPQWGLQWTEPESELPMPTPKIVLEPHIEVNSGKPNPCKASGWPPLPT 460
QY 380 NHEMLVAVAGIVLHPKIDFNHIDHFSVAIFTHKILPPDSQWVCSVNI VAGWVKPFI 439
Db 361 SEEMTLVRPDCGTVLQPNDFNHIDHFSVAIFTHKILPPDSQWVCSVNI VAGWVKPFI 420
QY 440 SVKVLKPLNAPNVTDGHNFAVINISSEPEFG-PKSCDKTHICP----- 484
Db 421 SVKVLKPLNAPNVTDGHNFAVINISSEPEFG-PKSCDKTHICP----- 480
QY 485 -----PCPAPELLGCP-----SVFLPPPKPKDTLMISRTPEV 516
Db 481 VTLNYLEPTDYELAVQIVRPPGEGEGHP:PVPRFTTASIGLPPRGLSLPKSQATALNL 540
QY 517 TCVVVDVSHEDPEVKFANWYVIGVVEVHNAKTKPPEQVNSTYRVVSVLT-VLHEDWLNCKE 575
Db 541 TWQPIETSSD-----EFY---VEVERWSQOTRSDQCN--IKVPGNLTSLVLLNLLFREQ 540
QY 576 YKCVSNKALPAPIKFTTSKAKQ:PPPEQVYT-----LPPSREEM-TKNQVSLTCLVK-- 527
Db 591 YSVRA-----PVNTKAG:FWSEETPAWTLSDLPDENIKITNITVYALVSWI 640
QY 628 ---GFYPSDIAVENESNCPENNY 648
Db 641 IVGYSSTSIIRYKVKQKNEQH 664

```

RESULT 9

US-08-484-438-10

: Sequence 10, Application US/08484448

: Patent No. 5811098

: Patent No. 5811098 5780031

: GENERAL INFORMATION:

: APPLICANT: Maisonnier, Gregory D.

: APPLICANT: Culouscou, Jean-Michel

: APPLICANT: Shoyab, Mohammed

: APPLICANT: Siedall, Clay B.

: APPLICANT: Hollistr m, Ingegerd

: APPLICANT: Hollistr m, Karl R.


```

: TITLE OF INVENTION: HER4 HUMAN RECEPTOR TYROSINE KINASE
: NUMBER OF SEQUENCES: 42
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Pennie & Edmonds
: STREET: 1155 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: U.S.A.
: ZIP: 10036-2711
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US 08/484,438
: FILING DATE: 07-JUN-1995
: CLASSIFICATION: 530
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/323,442
: FILING DATE: 14-OCT-1994
: APPLICATION NUMBER: US 08/150,704
: FILING DATE: 10-NOV-1993
: CLASSIFICATION: 530
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/981,165
: FILING DATE: 24-NOV-1992
: CLASSIFICATION: 530
: ATTORNEY/AGENT INFORMATION:
: NAME: Misrock, S. Leslie
: REGISTRATION NUMBER: 18,872
: REFERENCE/DOCKET NUMBER: 5624-230
: TELEPHONE: (212) 790 9090
: TELEFAX: (212) 869-8864/9741
: TELEX: 66141 PENNIE
: INFORMATION FOR SEQ ID NO: 10:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 911 amino acids
: TYPE: amino acid
: STRANDEDNESS: unknown
: TOPOLOGY: unknown
: MOLECULE TYPE: protein
: US-08-484-438-10

```

```

Query Match 33.9%; Score 1311; DB 2; Length 911;
Best Local Similarity 51.2%; Pred No 16045;
Matches 288; Conservative 27; Mismatches 64; Indels 184; Gaps 21.

QY 198 LFTSAFTRLIVR-----KCFACKWQHPGNHLETA--QMNNG-----VCHP 235
DB 477 LFTSINGRIVRNKRAENCTAI--QMVNHLGSSGCGGPGDQCLSCRRYSRGHCIE 534

QY 236 D TGE-----CI CPGFGMGRTCEK-----ACEHTFGRTIC 264
DB 535 SCNIYDGEFFRRFNGSLGVHQWQ-----QCKKMFVLI LCHQHPGPNQCKSHKIDPNC 589

QY 265 KENC SQQVCKSKYVVELHHPVQCSATQWKGLQCNKAECHPGHYGPKLPYS--QNNGE 321
DB 590 VEKQPCDQIGANSFTFKYALP-----DPEFP-----QHPN-TQGF-NFPI 629

QY 322 MCDRFQCCICSPQWGLQCHRPGLPRMTPKIYDLPIHLEVNSCKFPDCKASCWPLPTNF 381
DB 630 SHD-----CIYYP-WTC-----HSTLPQ-----HPKAVKA-----LECFP----- 658

QY 382 EMTLVKPDGTVLHPKDFNHTDHFSAVFTIHRILPDDSGWVCGSVNTVAGMVEKENISV 441
DB 659 --RLVGPD-----PFGCARPANTFLDPE----- 679

QY 442 KVLPKPLNAPNVIDTGHNAVINISSEPPYFGEKPKSCDKTHTCPCHPAPELLGQPSVFLFP 501
DB 660 -----EPKSCDKTHTCPCHPAPELLGQPSVFLFP 708

```

```

QY 502 PKPKDILMISPTPEVTCVVVEVSHEDPEVKFNWYVDRVEVHNAITKPPESGYNSTYKVV 561
DB 709 PKPKDILMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVVEVHAKIKPREQVNSTYKVV 768
QY 562 VLTVLHJLWLNWKEYKCKVSNKALPAPIEKTIISAKPQVFPFQVYTLFSPSEMTKRV 621
DB 769 VLTVLHJLWLNWKEYKCKVSNKALPAPIEKTIISAKQGFREFQVYTLFSPSEMTKRV 828
QY 622 LILVLKGYPSDIJAVWESNGQFNNNYKIHPPVLSDGSPFLYSKGLVDKSKWQGNVFS 681
DB 829 LILVLKGYPSDIJAVWESNGQFNNNYKIHPPVLSDGSPFLYSKGLVDKSKWQGNVFS 888
QY 682 CSVMHEALHNHYTQKSLSLSPCK 704
DB 889 CSVMHEALHNHYTQKSLSLSPCK 911

RESULT 10
US-08-385-329-4
: Sequence 4, Application US/08385229
: Patent No. 5605690
: GENERAL INFORMATION:
: APPLICANT: Jacobs, Cindy A.
: APPLICANT: Smith, Craig A.
: TITLE OF INVENTION: Method of Treating TNF-Dependent
: TITLE OF INVENTION: Inflammation Using Tumor Necrosis Factor Antagonists
: NUMBER OF SEQUENCES: 5
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Immunex Corporation
: STREET: 51 Duversity Street
: CITY: Seattle
: STATE: Washington
: COUNTRY: U.S.A.
: ZIP: 98101
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/385,229
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/07/946,236
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Wight, Christopher L.
: REGISTRATION NUMBER: 31,680
: REFERENCE/DOCKET NUMBER: 2503
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (206) 587-0430
: TELEFAX: (206) 587-0606
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 518 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-385-329-4

```

```

Query Match 33.6%; Score 1300; DB 1; Length 518;
Best Local Similarity 58.0%; Pred. No. 5,76-95;
Matches 285; Conservative 29; Mismatches 105; Indels 72; Gaps 17;

QY 248 GPTTEKAEELHTFGRTICPEKSGEGSKSYVFCLEPDPVQCSATQWKGLQCNKAECHPGFY 307
DB 66 GSTCHLRYHYDGTAGMCCSKCSPOGHAK--VPC-----TKISDVC-DSEENSTY 112
QY 308 GP-----DKLQKPSNNFE--MTTFEFG--GLTSPKRWGLQEFEGTTPPTPKIVD 354
US-08-385-329-4

```

```

; LENGTH: 486 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-243-010-1

Query Match
Rest: local similarity 33.68; Score 1298.5; DB 1; Length 486;
Matches 288; Conservative 26; Mismatches 104; Indels 71; Gaps 18;

QY 248 GRVTEKACELHIFGRICKERKSGGKSYVEKLPDPYGGSCATGKMGKLGQNEACHRGY 407
DB 37 GSTCELEVVYQIACMGCKSGKSPQIAK--VFG-----IKSDIVG--DSCELSTY 84
QY 308 GP-----DCKIRGSCNNGP--MCDPRQG--CLASPCWGLQTEREPIPRMTKIVD 454
DB 84 TQLMNVWPEVTSISGSPSSGQVETQA-TRFQNPDT-PPQWYVALSKFQGGPLAP-- 149
QY 355 LPDHIEVNSGRKNPICAKSGWPLPTNEEMTIV-KPDGTVLHPKDFNHTDHSVAITIHR 413
DB 140 -----LRKCRP---GFGVARPGTPTTSVVCKP-----CAPGTFSNTTS-STDIICRPHQ 183
QY 414 I-----LPDPSGV-WVGSVNT-----VAGMVKEKPFNISVK-----VLPKPLNAPNVIDTG 457
DB 184 IGVVAITPQNASMDAVTSTSTPTSPMAPAVHLPQPVSTESQHTQPTPESTAPST-- 249
QY 458 HNFVAVINISSHPYFC--EPKSKDHTICPPCPAPELGGHVSFLFPPKPKDILMISRTPE 515
DB 240 -SFLPMKPSPPAPFPEPKSCDKHTICPPCPAPELGGHVSFLFPPKPKDILMISRTPE 248
QY 516 VTCVVVDSHEDPEVKFNWYVDGVEVIINAKTKPREQYNSYRVSVLTIVLHQWLNKRE 575
DB 259 VTCVVVWVSHFETPKFNWYVDGVEVHNAKTKPREQYNSYRVSVLTIVLHQWLNKRE 457
QY 576 YKCKVSNKALPAIEKTIKSAKAGQPREPOVYTLPPSPPEMTKNQVSLTCLVKGFYPSDIA 635
DB 358 YKCKVSNKALPAIEKTIKSAKAGQPREPOVYTLPPSPPEMTKNQVSLTCLVKGFYPSDIA 417
QY 636 VFMKSNQGPENNYKTTPIDVDSGSPFLYSKLIIVDKSRWQGNVFSQSVMHREALHNYTG 695
DB 418 VEMESNGQPENNYKTTPFVLDSGSPFLYSKLIIVDKSRWQGNVFSQSVMHREALHNYTG 477
QY 696 KSLSLSPGK 704
DB 478 KSLSLSPGK 486

```

RESULT 12

```

US-08-784-512-3
; Sequence 3, Application US/08784512
; Patent No. 5872209
; GENERAL INFORMATION:
; APPLICANT: BARNIK, Eckart
; APPLICANT: EIDENMUELLER, Bernd
; APPLICANT: BUETTNER, Frank
; APPLICANT: CATHERSON, Bruce
; APPLICANT: HUGHES, Clare
; TITLE OF INVENTION: An artificial recombinant substrate (RAGG 1)
; TITLE OF INVENTION: and native aggrecan to study the proteolytic activity of
; TITLE OF INVENTION: "Aggrecanase" in cell culture systems
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: Suite 500, 3000 K Street, N.W.
; CITY: Washington, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; SEQUENCE CHARACTERISTICS:

RESULT 11
US-08-243-010-1
; Sequence 1, Application US/08243010
; Patent No. 5639597
; GENERAL INFORMATION:
; APPLICANT: Laulier, Leander
; APPLICANT: ZellmiesseL, Gerd
; APPLICANT: Aquendo, Patricia
; TITLE OF INVENTION: Cell free Receptor Binding Assays, the
; TITLE OF INVENTION: Production and Use Thereof
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Elinegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dumer
; STREET: 1400 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3415
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/243,010
; FILING DATE: 13-MAY-1994
; CLASSIFICATION: 415
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/798,564
; FILING DATE: 26-NOV-1991
; APPLICATION NUMBER: DE P 40 37 817.3
; FILING DATE: 28-NOV-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Einadl, Carol P.
; REGISTRATION NUMBER: 32,220
; REFERENCE/ID#&RET INFO#P# 024W1-1132-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:

```

: CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/784,512
 : FILING DATE: 17-JAN-1997
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: EP 96100682.2
 : FILING DATE: 18-JAN-1996
 : ATTORNEY/AGENT INFORMATION:
 : NAME: GRANADOS, Patricia D.
 : REGISTRATION NUMBER: 33,683
 : REFERENCE/DOCKET NUMBER: 18748/311
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (202)572-5300
 : TELEFAX: (202)672-5399
 : TELEX: 904136
 : INFORMATION FOR SEQ ID NO: 3:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 396 amino acids
 : TYPE: amino acid
 : STRANDEDNESS: single
 : TOPOLOGY: linear
 : MOLECULE TYPE: protein
 : FEATURE:
 : NAME/KEY: Protein
 : LOCATION: 1..396
 : US-08-784-512-3

Query Match 33.1% Score 1279.5; DB 2; Length 396;
 Best Local Similarity 68.4%; Pred. No. 1.7e-93;
 Matches 255; Conservative 21; Mismatches 62; Indels 35; Gaps 1;
 QY 353 VDLPHIEVNSKFNPKICKASGP---LPTNEEMTLVKPDGTVL-----HPKDFNH 400
 Db 38 VDIPEFEGVGGEEDITVQTVTPDMELPLRNITEGEARGSVILTVKPIFVSPSLEP 97
 QY 401 IDHFSVAIFTIHRILPPDSGVWV-CSVNTVAGMVKPKPNI SVKVLKPKLNAPNVIDTGNH 459
 Db 98 EEPFTFA-----PEIGATAFAVENETGEATPFWG-----FTTGLGPATAFTSED 143
 QY 460 FA--VINISSEPF-----GPKSKDTHGCPGAPMLGQPSVFLPPKPKDTLMS 511
 Db 144 LVQVTAVPQGPPLPGGDPPEPKSCUKTHGCPGAPMLGQPSVFLPPKPKDTLMS 203
 QY 512 RPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWL 571
 Db 204 RTPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWL 263
 QY 572 NGKEYCKVSNKALPAPIEKTIISKAKQPRPQVYTLPPSRPEMTKNQVSLTCLVKGFP 631
 Db 264 NGKEYCKVSNKALPAPIEKTIISKAKQPRPQVYTLPPSRPEMTKNQVSLTCLVKGFP 323
 QY 632 SDIAVWESNGQPNYKTTTPVLDSDGSPFLYSKLIVKSKWQOCNVFSCSVMHHEALHN 691
 Db 324 SDIAVWESNGQPNYKTTTPVLDSDGSPFLYSKLIVKSKWQOCNVFSCSVMHHEALHN 383
 QY 692 HYTKSLSLSPGK 704
 Db 384 HYTKSLSLSPGK 396
 RESULT 13
 US-09-176-228-3
 : Sequence 3, Application US/09176228
 : Patent No. 618034
 : GENERAL INFORMATION:
 : APPLICANT: RABINIK, Eckart
 : APPLICANT: EIDENMUELLER, Bernd
 : APPLICANT: BUETTINGER, Frank
 : APPLICANT: CATERSON, Bruce
 : APPLICANT: HUGHES, Clare
 : TITLE OF INVENTION: An artificial recombinant substrate (rAGG 1)
 : TITLE OF INVENTION: And native aggracan to study the proteolytic activity of
 : TITLE OF INVENTION: "Aggracanase" in cell culture systems

: NUMBER OF SEQUENCES: 4
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Foley & Lardner
 : STREET: Suite 500, 3000 K Street, N.W.
 : CITY: Washington, D.C.
 : COUNTRY: USA
 : ZIP: 20007-5109
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Floppy disk
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: PatentIn Release #1.0, Version #1.25
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/09/176,228
 : FILING DATE:
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: US/08/784,512
 : FILING DATE: 17-JAN-1997
 : APPLICATION NUMBER: EP 96100682.2
 : FILING DATE: 18-JAN-1996
 : ATTORNEY/AGENT INFORMATION:
 : NAME: GRANADOS, Patricia D.
 : REGISTRATION NUMBER: 33,683
 : REFERENCE/DOCKET NUMBER: 18748/311
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (202)672-5300
 : TELEFAX: (202)672-5399
 : TELEX: 904136
 : INFORMATION FOR SEQ ID NO: 3:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 396 amino acids
 : TYPE: amino acid
 : STRANDEDNESS: single
 : TOPOLOGY: linear
 : MOLECULE TYPE: protein
 : FEATURE:
 : NAME/KEY: Protein
 : LOCATION: 1..396
 : US-09-176-228-3

Query Match 33.1% Score 1279.5; DB 4; Length 396;
 Best Local Similarity 68.4%; Pred. No. 1.7e-93;
 Matches 255; Conservative 21; Mismatches 62; Indels 35; Gaps 7;
 QY 353 VDLPHIEVNSKFNPKICKASGP---LPTNEEMTLVKPDGTVL-----HPKDFNH 400
 Db 38 VDIPEFEGVGGEEDITVQTVTPDMELPLRNITEGEARGSVILTVKPIFVSPSLEP 97
 QY 401 IDHFSVAIFTIHRILPPDSGVWV-CSVNTVAGMVKPKPNI SVKVLKPKLNAPNVIDTGNH 459
 Db 98 EEPFTFA-----PEIGATAFAVENETGEATPFWG-----FTTGLGPATAFTSED 143
 QY 460 FA--VINISSEPF-----GPKSKDTHGCPGAPMLGQPSVFLPPKPKDTLMS 511
 Db 144 LVQVTAVPQGPPLPGGDPPEPKSCUKTHGCPGAPMLGQPSVFLPPKPKDTLMS 203
 QY 512 RPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWL 571
 Db 204 RTPEVTCVVVDVSHEDPEVKFNMYVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWL 263
 QY 572 NGKEYCKVSNKALPAPIEKTIISKAKQPRPQVYTLPPSRPEMTKNQVSLTCLVKGFP 631
 Db 264 NGKEYCKVSNKALPAPIEKTIISKAKQPRPQVYTLPPSRPEMTKNQVSLTCLVKGFP 323
 QY 632 SDIAVWESNGQPNYKTTTPVLDSDGSPFLYSKLIVKSKWQOCNVFSCSVMHHEALHN 691
 Db 324 SDIAVWESNGQPNYKTTTPVLDSDGSPFLYSKLIVKSKWQOCNVFSCSVMHHEALHN 383
 QY 692 HYTKSLSLSPGK 704
 Db 384 HYTKSLSLSPGK 396

```

Query Match      33.0%  Score 1276.5;  DB 1;  length 438;
Host Local Similarity 55.5%  Pred. No. 3,3e-93;
Matches 274;  Conservative 23;  Mismatched 86;  Indels 109;  Gaps 14;

QY      247  MGRITCEKA  CELHIF--GRTCKEPCSGEGSGKSVVF--LPPYGSAT--AWKLTG--NEA-- 301
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      18  LGVTAARRLNCKWKHIYPSGHCCKRQCHGDMVN-----RCCHTRD 57

QY      302  --CHP--GPGY-----DPRKTLKPSGNGGHCDFGQCL-----GSPWGQGLQCHRE 343
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      58  TLCHPCETGFEYNAVNYDLKJLQJLNIURSSSELKUNLPIQJTVCKRCRPGTQ----- 110

QY      344  GIPRMTPKLVDPDRIEVSCKNP-----ICKASGWL--PINEMTIAVKPGDT 391
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      111  --PRDSQKGLVPCVCPDCHGSPGNNQACKWEINELISGQJLPHPASLILVACHERS 168

QY      392  VLIHKPDENHICHESVAIFTHPIIPPSASVWVSNTVAGSMVEKPEFI--SVFVLPKPLNA 450
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      169  LL-----ALLWELQRIPI-----RPIVQSTIVVWPIPSFL 199

QY      451  PKVDTGHNFAVNISSHPYFGHPKSCDKTHYCPGCPAPELGGHVSFLPPKPKDTLMJ 510
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      200  P-----STETLVEPS--TKTHF--P--PAPEAESASVFLPEPKPKDTLMJ 244

QY      511  SKPEPVTGVVIVSHSPGPKFNMYVLCGVVHNAKTKPREGQYNSIYRVSVSLVILVHDW 570
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      245  SKPEPVTGVVIVSHSPGPEVFNMYVLCGVVHNAKTKPREGQYNSIYRVSVSLVILVHDW 604

QY      571  LMGKEYKCKVSNKALPAPIERTKSKAGGAPPEPQVITLPPSPPEEMTKKCVSLTLVKGFY 630

```

```

QY 392 VLHPKDFMHTHPSVALETHIRILPDSGVWVGSVNTVAGMVEKPENI-SVKVLPKPLNA 450
Db 169 LL-----ATLWMTQRIPTF-----RPTTVOSTTIVWPTSMI 199
QY 451 PNVIDTGHNEAVINISSEPYGEPKSCDKTHTCPPCPAPELGGPSVFLFPKPKDILMI 510
Db 200 P-----STPTLVEFRSCDKTHTCPPCPAPEAEGAPSVFLFPKPKDILMI 244
QY 511 SRTPETVCVVVDVSHHDPVEVKFNWYVDGVGVHNAKIKPREEOYNSTYRVVSVLTIVLHODW 570
Db 245 SRTPETVCVVVDVSHHDPVEVKFNWYVDGVGVHNAKIKPREEOYNSTYRVVSVLTIVLHODW 304
QY 571 LNKKEYCKVSNKALFAPLEKTTISKAKGQPREPQVYITLPPSREEMTKNAVSLTCLVKGY 630
Db 305 INCKEYCKVSNKALFAPLEKTTISKAKGQPREPQVYITLPPSREEMTKNAVSLTCLVKGY 364
QY 631 PSDIAVEWESNGOPENNYKTTTPVLDSDGSFFLYSKLTVDKSWOOGNVFSCSVMEALH 690
Db 365 PSDIAVEWESNGOPENNYKTTTPVLDSDGSFFLYSKLTVDKSWOOGNVFSCSVMEALH 424
QY 691 NHYTKSLSLSPGK 704
Db 425 NHYTKSLSLSPGK 438

```

Search completed: September 23, 2002, 22:10:22
Job time: 423 sec

100

101

102

103

104

105

106

107

108

109

110

111

112

113

114

115

116

117

118

119

120

121

122

123

124

125

126

127

128

129

130

131

132

133

134

135

136

137

138

139

140

141

142

143

144

145

146

147

148

149

150

151

152

153

154

155

156

157

158

159

160

161

162

163

164

165

166

167

168

169

170

171

172

173

174

175

176

177

178

179

180

181

182

183

184

185

186

187

188

189

190

191

192

193

194

195

196

197

198

199

200

201

202

203

204

205

206

207

208

209

210

211

212

213

214

215

216

217

218

219

220

221

222

223

224

225

226

227

228

229

230

231

232

233

234

235

236

237

238

239

240

241

242

243

244

245

246

247

248

249

250

251

252

253

254

255

256

257

258

259

260

261

262

263

264

265

266

267

268

269

270

271

272

273

274

275

276

277

278

279

280

281

282

283

284

285

286

287

288

289

290

291

292

293

294

295

296

297

298

299

300

301

302

303

304

305

306

307

308

309

310

311

312

313

314

315

316

317

318

319

320

321

322

323

324

325

326

327

328

329

330

331

332

333

334

335

336

337

338

339

340

341

342

343

344

345

346

347

348

349

350

351

352

353

354

355

356

357

358

359

360

361

362

363

364

365

366

367

368

369

370

371

372

373

374

375

376

377

378

379

380

381

382

383

384

385

386

387

388

389

390

391

392

393

394

395

396

397

398

399

400

401

402

403

404

405

406

407

408

409

410

411

412

413

414

415

416

417

418

419

420

421

422

423

424

425

426

427

428

429

430

431

432

433

434

435

436

437

438

439

440

441

442

443

444

445

446

447

448

449

450

451

452

453

454

455

456

457

458

459

460

461

462

463

464

465

466

467

468

469

470

471

472

473

474

475

476

477

478

479

480

481

482

483

484

485

486

487

488

489

490

491

492

493

494

495

496

497

498

499

500

501

502

503

504

505

506

507

508

509

510

511

512

513

514

515

516

517

518

519

520

521

522

523

524

525

526

527

528

529

530

531

532

533

534

535

536

537

538

539

540

541

542

543

544

545

546

547

548

549

550

551

552

553

554

555

556

557

558

559

560

561

562

563

564

565

566

567

568

569

570

571

572

573

574

575

576

577

578

579

580

581

582

583

584

585

586

587

588

589

590

591

592

593

594

595

596

597

598

599

600

601

602

603

604

605

606

607

608

609

610

611

612

613

614

615

616

617

618

619

620

621

622

623

624

625

626

627

628

629

630

631

632

633

634

635

636

637

638

639

640

641

642

643

644

645

646

647

648

649

650

651

652

653

654

655

656

657

658

659

660

661

662

663

664

665

666

667

668

669

670

671

672

673

674

675

676

677

678

679

680

681

682

683

684

685

686

687

688

689

690

691

692

693

694

695

696

697

698

699

700

701

702

703

704

705

706

707

708

709

710

711

712

713

714

715

716

717

718

719

720

721

722

723

724

725

726

727

728

729

730

731

732

733

734

735

736

737

738

739

740

741

742

743

744

745

746

747

748

749

750

751

752

753

754

755

756

757

758

759

760

761

762

763

764

765

766

767

768

769

770

771

772

773

774

775

776

777

778

779

780

781

782

783

784

785

786

787

788

789

790

791

792

793

794

795

796

797

798

799

800

801

802

803

804

805

806

807

808

809

810

811

812

813

814

815

816

817

818

819

820

821

822

823

824

825

826

827

828

829

830

831

832

833

834

835

836

837

838

839

840

841

842

843

844

845

846

847

848

849

850

851

852

853

854

855

856

857

858

859

860

861

862

863

864

865

866

867

868

869

870

871

872

873

874

875

876

877

878

879

880

881

882

883

884

885

886

887

888

889

890

891

892

893

894

895

896

897

898

899

900

901

902

903

904

905

906

907

908

909

910

911

912

913

914

915

916

917

918

919

920

921

922

923

924

925

926

927

928

929

930

931

932

933

934

935

936

937

938

939

940

941

942

943

944

945

946

947

948

949

950

951

952

953

954

955

956

957

958

959

960

961

962

963

964

965

966

967

968

969

970

971

972

973

974

975

976

977

978

979

980

981

982

983

984

985

986

987

988

989

990

991

992

993

994

995

996

997

998

999

1000

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd

OM protein - protein search, using sw model

Run on: September 23, 2002, 22:08:44, Search time: 45.53 Seconds
(without alignments)
1485.765 Million cell updates/sec

Title: US-09-713 764.2
 perfect score: 3868
 Sequence: 1 MRSIASIVVNGSTIIISIV
 Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 36089334 residues
Total number of hits satisfying chosen parameters: 283138

```

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing. Minimum Match: 92
Maximum Match: 100
Listing first 45

```

```
Database :   p1r_71:*
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*
```

SUMMARIES

Result No.	Query %		Length	DE	ID	Description
	Match	Score				
1	25.3	67.8	1126	1	558388	protein-tyrosine k
2	24.58	63.5	1126	1	557845	protein-tyrosine k
3	24.35	62.9	1126	2	754237	protein-tyrosine k
4	24.35	62.9	1123	1	UN0712	protein-tyrosine k
5	23.63	61.1	1125	1	460771	protein-tyrosine k
6	12.71	32.9	374	2	554339	lg heavy chain v
7	12.62	32.5	370	1	GHRP	lg gamma-1 chain c
8	12.49	32.3	355	4	S31866	lg gamma-1 chain c
9	12.04	31.1	234	2	PRO207	lg gamma-1 chain c
10	11.83	30.4	372	2	A23511	lg gamma-3 chain c
11	11.81	30.5	372	2	A40764	lg gamma-3 chain c
12	11.58	29.9	289	1	C3H7W1	lg gamma-3 heavy c
13	11.54	29.8	356	1	22HR	lg gamma-2 chain c
14	11.44	29.6	357	1	G4HD	lg gamma-2 chain c
15	9.32	24.2	323	1	GHRB	lg gamma-2 chain c
16	9.15	23.2	328	2	147159	lg gamma-2a chain
17	9.10.5	23.5	328	2	147160	lg gamma-2a chain
18	9.06.5	23.4	272	2	147162	lg gamma-2b chain c
19	8.93.5	23.1	329	1	G20F	lg gamma-2 chain c
20	8.92	23.1	328	2	147158	lg gamma-2 chain c
21	8.83	22.8	328	2	147161	lg gamma-3 chain c
22	8.82.5	22.8	1136	1	S57845	protein-tyrosine k
23	8.74.5	22.6	1138	1	S21066	protein-tyrosine k
24	8.66	22.4	470	2	S25080	lg heavy chain pre
25	8.57.5	22.2	325	1	GMS2	lg gamma-3 chain c
26	8.51.5	22.0	333	2	FS0018	lg gamma-2b chain
27	8.49.5	22.0	308	2	360554	lg heavy chain c r
28	8.49.5	22.0	472	2	S31459	lg gamma-1 chain
29	8.48	21.9	444	2	PC4433	monoclonal antibody

ALIGNMENTS

RESULT 1
158386
protein tyrosine kinase (BC 2.7.1.112), receptor type tek precursor - human
S:Species: Homo sapiens (man)
C:Date: 02-Jul-1997 #sequence_revision 02 Jul-1997 #text_change 16-Jul-1999
C:Accession: I58388
R:Ziegler, S.F.; Bird, T.A.; Schneider, J.A.; Schooley, K.A.; Baum, P.R.
Oncogene 8, 663-670, 1993
A:Title: Molecular cloning and characterization of a novel receptor protein tyrosine
A:Reference number: 158388; MUID:93173509
A:Accession: 158388
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1124 <RES>
A:Cross references: GB:L6133; NID:g226823, FIDN.AAAG1139.1, PID:g23824
C:Genetics:
A:Gene: GDB:TEK
A:Cross-references: GDB:344185; OMIM:600221
A:Map position: 9p21-9p21
C:Function:

```

121 QASFLPATITMTVDKGNWNISFKVLLKEEPAVLYKNGSETHSVPRHEVDILEVHLPH 180
|||||
121 QASFLPATITMTVDKGNWNISFKVLLKEEPAVLYKNGSETHSVPRHEVDILEVHLPH 180
|||||
181 AOHQAGVYSARYIGUNLFTSAFTPLIVPPCEAKWPEKPEKTHLCAAMNNVWHEDETCG 240
|||||
181 AOHQAGVYSARYIGUNLFTSAFTPLIVPPCEAKWPEKPEKTHLCAAMNNVWHEDETCG 240
|||||
241 ICPPGEMGTCEKAKVELHITGRTIKERSSQFCKKSYVFLPLDPPYYSQATGSKGLQNE 300
|||||
241 ICPPGEMGTCEKAKVELHITGRTIKERSSQFCKKSYVFLPLDPPYYSQATGSKGLQNE 300
|||||
301 ACHPGYCHDCKLRSSNGGEMGRTPLDAPSPWGLQCHREGIPRMPTKIVLPLDPIIE 360
|||||
301 ACHPGYCHDCKLRSSNGGEMGRTPLDAPSPWGLQCHREGIPRMPTKIVLPLDPIIE 360
|||||
361 VNSGKFNPTCKASGMWLPENHMTLVKQGVIVLHPKDNHHDHFSVAFTIHRILPPDSG 420
|||||
361 VNSGKFNPTCKASGMWLPENHMTLVKQGVIVLHPKDNHHDHFSVAFTIHRILPPDSG 420
|||||
421 VVWCSVNTVAGWVEKPEKNSVAVLPEKFNAPRVDTGHNFAVINISSEPVGEKSCDKT 480
|||||
421 VVWCSVNTVAGWVEKPEKNSVAVLPEKFNAPRVDTGHNFAVINISSEPVGEKSCDKT 480
|||||
473
481 HTCPGCPAPELIGSGSVFLFPPEKPKDTLMSIKTEPVTCVVDVSHDEPEVKNNYV 536
|||||
481 HTCPGCPAPELIGSGSVFLFPPEKPKDTLMSIKTEPVTCVVDVSHDEPEVKNNYV 536
|||||
536 KAKGCPREPQVYTLPESEKEMTKNVSSTCLVKGKFPYS---DIAVEWESNGUPENN 647
|||||
536 KAKGCPREPQVYTLPESEKEMTKNVSSTCLVKGKFPYS---DIAVEWESNGUPENN 647
|||||
539 ASIG-----LPPKCLSLILPKSQTLLNLTWQPIFPSSDDPYVEVERKSVQMS 587
|||||
539 ASIG-----LPPKCLSLILPKSQTLLNLTWQPIFPSSDDPYVEVERKSVQMS 587
|||||

RESULT 3
154237
protein-tyrosine kinase (EC 2.7.1.112), receptor type tek precursor mouse
N.Alternate names: protein-tyrosine kinase receptor type tie2
C.Species: Mus sp. (mouse)
C.Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 21-Jan-2000
C.Accession: 154237
R.Kuntz, A.S., Stackel, S.A., Wilks, A.F.
Growth Factors 9, 99-105, 1993
A.Title: tie2, a putative protein tyrosine kinase from a new class of cell surface re
A.Reference number: 154237; MUID:9403116
A.Accession: 154237
A.Status: preliminary, translated from cDNA, EMBL, DDBJ
A.Molecule type: mRNA
A.Residues: 1-1122 <RES>

F:542-626/Domain: fibronectin type III repeat homology <FN3B>
F:639-743/Domain: fibronectin type III repeat homology <FN3C>
F:753-773/Domain: transmembrane #status predicted <TM>
F:823-1130/Domain: protein kinase #status predicted <TK>
F:831-839/Region: protein kinase ATP-binding motif
F:140,169,199,499,444,490,497,450,492/Binding site: catalytic site (Asn) (catalytic) #st
F:856,875,965/Active site: Lys, Glu, Asp #status predicted

Query Match 63.5%; Score 2458; DB 1; Length 1125;
Best Local Similarity 71.6%; Pred. No. 3.6e-140;
Matches 471; Conservative 35; Mismatches 76; Indels 82; Gaps 7;

QY 1 MSLASLVAGVSVLLSGIVGAMILLINSLPVSVAETSLICIASGRHHEPITIGRD 60
|||||
DB 1 MSLASLVAGVSVLLSATVAGAMILLINSLPVSVAETSLICIASGRHHEPITIGRD 60
|||||
QY 61 FPA: MNCQDPI FVGVWVTPWAKKVVWFKASKINZAYDEGVPVEEALPIPIKMPQ 120
|||||
DB 61 FPA: MNCQDPI FVGVWVTPWAKKVVWFKASKINZAYDEGVPVEEALPIPIKMPQ 120
|||||
QY 121 QASFLPATITMTVDKGNWNISFKVLLKEEPAVLYKNGSETHSVPRHEVDILEVHLPH 180
|||||
DB 121 QASFLPATITMTVDKGNWNISFKVLLKEEPAVLYKNGSETHSVPRHEVDILEVHLPH 180
|||||
QY 181 AOHQAGVYSARYIGUNLFTSAFTPLIVPPCEAKWPEKPEKTHLCAAMNNVWHEDETCG 240
|||||
DB 181 AOHQAGVYSARYIGUNLFTSAFTPLIVPPCEAKWPEKPEKTHLCAAMNNVWHEDETCG 240
|||||
QY 241 ICPPGEMGTCEKAKVELHITGRTIKERSSQFCKKSYVFLPLDPPYYSQATGSKGLQNE 300
|||||
DB 241 ICPPGEMGTCEKAKVELHITGRTIKERSSQFCKKSYVFLPLDPPYYSQATGSKGLQNE 300
|||||
QY 301 ACHPGYCHDCKLRSSNGGEMGRTPLDAPSPWGLQCHREGIPRMPTKIVLPLDPIIE 360
|||||
DB 301 ACHPGYCHDCKLRSSNGGEMGRTPLDAPSPWGLQCHREGIPRMPTKIVLPLDPIIE 360
|||||
QY 361 VNSGKFNPTCKASGMWLPENHMTLVKQGVIVLHPKDNHHDHFSVAFTIHRILPPDSG 420
|||||
DB 361 VNSGKFNPTCKASGMWLPENHMTLVKQGVIVLHPKDNHHDHFSVAFTIHRILPPDSG 420
|||||
QY 421 VVWCSVNTVAGWVEKPEKNSVAVLPEKFNAPRVDTGHNFAVINISSEPVGEKSCDKT 480
|||||
DB 421 VVWCSVNTVAGWVEKPEKNSVAVLPEKFNAPRVDTGHNFAVINISSEPVGEKSCDKT 480
|||||
QY 481 HTCPGCPAPELIGSGSVFLFPPEKPKDTLMSIKTEPVTCVVDVSHDEPEVKNNYV 536
|||||
DB 481 HTCPGCPAPELIGSGSVFLFPPEKPKDTLMSIKTEPVTCVVDVSHDEPEVKNNYV 536
|||||
QY 536 KAKGCPREPQVYTLPESEKEMTKNVSSTCLVKGKFPYS---DIAVEWESNGUPENN 647
|||||
DB 536 KAKGCPREPQVYTLPESEKEMTKNVSSTCLVKGKFPYS---DIAVEWESNGUPENN 647
|||||
QY 539 ASIG-----LPPKCLSLILPKSQTLLNLTWQPIFPSSDDPYVEVERKSVQMS 587
|||||
DB 539 ASIG-----LPPKCLSLILPKSQTLLNLTWQPIFPSSDDPYVEVERKSVQMS 587
|||||

RESULT 3
154237
protein-tyrosine kinase (EC 2.7.1.112), receptor type tek precursor mouse
N.Alternate names: protein-tyrosine kinase receptor type tie2
C.Species: Mus sp. (mouse)
C.Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 21-Jan-2000
C.Accession: 154237
R.Kuntz, A.S., Stackel, S.A., Wilks, A.F.
Growth Factors 9, 99-105, 1993
A.Title: tie2, a putative protein tyrosine kinase from a new class of cell surface re
A.Reference number: 154237; MUID:9403116
A.Accession: 154237
A.Status: preliminary, translated from cDNA, EMBL, DDBJ
A.Molecule type: mRNA
A.Residues: 1-1122 <RES>

```


A:Cross-references: GR:567951; NID:q452873; EMDN:AAH28663; E:1004450474

C:Genetics:

A:Gene: tie2

C:Superfamily: protein-tyrosine kinase, receptor type III, EGF homology, fibronectin type III

C:Keywords: ATP, autophosphorylation, duplication, glycoprotein, phosphoprotein, phospho

F:337-104/Domain: immunoglobulin homology <IMM>

F:211-251/Domain: EGF homology <EGF>

F:255-248/Domain: EGF homology <EGF>

F:302-340/Domain: fibronectin type III repeat homology <FN3>

F:540-624/Domain: fibronectin type III repeat homology <FN3>

F:820-1097/Domain: protein kinase homology <PK>

F:828-836/Region: protein kinase ATP-binding motif

Query Match 62.9%; Score 2432.5; DB 1; Length 1123;
Best Local Similarity 68.1%; Pred. No. 9, 6e-129;
Matches 481; Conservative 50; Mismatches 94; Indels 81; Gaps 14.

```

QY 1 MDSIASIVLCGVSLISGVTEGAMDLILNSLPLVSDAETSLTICIASGWPHPEPTIGRD 60
DB 1 MDSLAGLVLCGVSLLYGVVEGAMDLILNSLPLVSDAETSLTICIASGWPHPEPTIGRD 60
QY 61 FEALMNHQDPLEVTQVTPEWAKVVKPEKASKINGAYFCGRVPGEAIPRTMKMPQ 120
DB 61 FEALMNHQDPLEVTQVTPEWAKVVKPEKASKINGAYFCGRVPGEAIPRTMKMPQ 120
QY 121 QASFLPATLMTVDKGNVNSIFSKVLIKEEDAVIYKNGSFIHSVPHEVEDILEVHLPH 180
DB 121 QASFLPATLMTVDKGNVNSIFSKVLIKEEDAVIYKNGSFIHSVPHEVEDILEVHLPH 180
QY 181 AQPDAGVYSARYIGGNLFSAFTKILVRKCEAKWKPCPNHCTACMNNGVCHEDIGEC 240
DB 181 AQPDAGVYSARYIGGNLFSAFTKILVRKCEAKWKPCPNHCTACMNNGVCHEDIGEC 240
QY 241 ICPPGEMKICEKACELHTEFRICKERKSAQEGKSKVPELPPPYGSAFENKGLQANE 300
DB 241 ICPPGEMKICEKACELHTEFRICKERKSAQEGKSKVPELPPPYGSAFENKGLQANE 300
QY 301 AHPHFVCEPKLEPSNNEMDPEFACILSPKQWGLQERESLPHMTKPIVGLFGLIE 360
DB 301 AHPHFVCEPKLEPSNNEMDPEFACILSPKQWGLQERESLPHMTKPIVGLFGLIE 360
QY 361 VNSKFNPTCKASGWLPTNEMILVLPKDGIVLHPKDFNHTDHFSALEPIHKLDPDSC 420
DB 361 VNSKFNPTCKASGWLPTNEMILVLPKDGIVLHPKDFNHTDHFSALEPIHKLDPDSC 420
QY 421 VVVVSVNTVAGWKEKPNISVKKVLPKPLNPNVITDGNFAVINISSEFYFS 472
DB 421 VVVVSVNTVAGWKEKPNISVKKVLPKPLNPNVITDGNFAVINISSEFYFS 472
QY 473 VVSGSVNTVAGWKEKPNISVKKVLPKPLNPNVITDGNFAVINISSEFYFS 496
DB 473 VVSGSVNTVAGWKEKPNISVKKVLPKPLNPNVITDGNFAVINISSEFYFS 496
QY 481 LFYKPVQCAWKYIEVTNEITLNYLEPTDYELC--VQLAPPDGRGRSHGPPVPTTAS 538
DB 481 LFYKPVQCAWKYIEVTNEITLNYLEPTDYELC--VQLAPPDGRGRSHGPPVPTTAS 538
QY 497 VHLFPPKPKDTLMSRTP-FVTCVVVDVSHEDPEVKFNWYVDGVGVVHNAKTKPREQYNS 555
DB 497 VHLFPPKPKDTLMSRTP-FVTCVVVDVSHEDPEVKFNWYVDGVGVVHNAKTKPREQYNS 555
QY 539 IGLPPLPRLSLPSQTALNTLTPQITNSD-----EFY---VEVERPSLQTTSDQDN 589
DB 539 IGLPPLPRLSLPSQTALNTLTPQITNSD-----EFY---VEVERPSLQTTSDQDN 589
QY 556 TYRVSVLT-VLHGDWLNKREYKCKVSNKALPALEKTSIKAKQV-PREPVVT----LP 609
DB 556 TYRVSVLT-VLHGDWLNKREYKCKVSNKALPALEKTSIKAKQV-PREPVVT----LP 609
QY 590 -IKVFGNLTSLVLSNIVTRPGVTVRA-----RVNTKAQGRWSEERLAWTSLDLP 638
DB 590 -IKVFGNLTSLVLSNIVTRPGVTVRA-----RVNTKAQGRWSEERLAWTSLDLP 638
QY 610 PSREEM-----TKNVSSTLCVKGFPSPDIAVWESNCGPENNY 648
DB 610 PSREEM-----TKNVSSTLCVKGFPSPDIAVWESNCGPENNY 648
QY 639 PQPENIKISNITDSTAMVSWT-IVDGSYSISIIIRYKVGQKNEQD 684
DB 639 PQPENIKISNITDSTAMVSWT-IVDGSYSISIIIRYKVGQKNEQD 684

```

RESULT 4

JN0712

N:protein-tyrosine kinase (EC 2.7.1.112), receptor type tek precursor mouse

C:Species: Mus musculus (house mouse)

C:Date: 27-Jun-1994 #sequence_revision 27 Jun 1994 #text_change 16 Jul-1999

C:Accession: JN0712; S57848, S43495, S43494, S33142
R:Watanabe, A.; Hamaguchi, I.; Hashiyama, M.; Murayama, Y.; Yasunaga, K.; Suda, T.
Biochem. Biophys. Res. Commun. 195, 301-309, 1993
A:Title: Molecular cloning and characterization of mouse tie-2 and IER receptor tyrosin
A:Reference number: JN0712; MUID:93371421
A:Accession: JN0712
A:Molecule type: mRNA
A:Residues: 1-1123 <IMA>
F:Satoh, T.N.; Qiu, Y.; Kozak, C.A.; Andus, K.L.
Proc. Natl. Acad. Sci. U.S.A. 90, 9355-9358, 1993
A:Title: tie-2 and tie-2 define another class of putative receptor tyrosine kinase ge
A:Reference number: S57845; MUID:94023374
A:Accession: S57848
A:Status: preliminary, nucleic acid sequence not shown; translation not shown
A:Molecule type: mRNA
A:Residues: 1-766 <IMM>
A:Cross-references: EMBL:X71426, NID:4296612, EMBL:U0557.1, FID:4296613
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1993
R:Kumamoto, D.J.
submitted to the EMBL Data Library, April 1993
A:Reference number: S43495
A:Accession: S43495
A:Molecule type: mRNA
A:Residues: 1-537 <IMM>
A:Cross-references: EMBL:X71426, NID:4296612, EMBL:U0557.1, FID:4296613
R:Kumamoto, D.J.; Yamaguchi, T.P.; Conlon, R.A.; Rossant, J.; Breitman, M.L.
Oncogene 7, 1471-1480, 1992
A:Title: tek, a novel tyrosine kinase gene located on mouse chromosome 4, is expresse
A:Reference number: S43494; MUID:92334855
A:Accession: S43494
A:Molecule type: mRNA
A:Residues: 823-1123 <IMM>
A:Cross-references: EMBL:X57553
C:Comment: Receptor tyrosine kinase-ligand systems play an important role in the cons

C:Genetics:

A:Map position: 4

C:Superfamily: protein-tyrosine kinase, receptor type III, EGF homology, fibronectin
C:Keywords: ATP, autophosphorylation, duplication, glycoprotein, phosphoprotein, phos
F:1-25/Domain: signal sequence #status predicted <SIG>
F:23-1123/Region: protein tyrosine kinase, receptor type tie-2 #status predicted <MA

F:37-104/Domain: immunoglobulin homology <IM1>

F:135-137/Region: cell attachment (R-G-D) motif

F:211-251/Domain: EGF homology <EGF>

F:255-248/Domain: EGF homology <EGF>

F:302-340/Domain: EGF homology <EGF>

F:364-426/Domain: immunoglobulin homology <IM2>

F:447-525/Domain: fibronectin type III repeat homology <FN3A>

F:540-624/Domain: fibronectin type III repeat homology <FN3B>

F:637-719/Domain: fibronectin type III repeat homology <FN3C>

F:751-773/Domain: transmembrane #status predicted <TM>

F:821-1098/Region: protein kinase homology <PK>

F:829-837/Region: protein kinase ATP-binding motif

F:140-158,390,438,464,558,595,648,690/Binding site: carbohydrate (Asn) (covalent) #st

F:854,871,963/Active site: Lys, Glu, Asp #status predicted

Query Match 62.9%; Score 2432.5; DB 1; Length 1123;

Best Local Similarity 68.1%; Pred. No. 9, 6e-129;

Matches 481; Conservative 50; Mismatches 94; Indels 81; Gaps 14;

```

QY 1 MDSIASIVLCGVSLISGVTEGAMDLILNSLPLVSDAETSLTICIASGWPHPEPTIGRD 60
DB 1 MDSLAGLVLCGVSLLYGVVEGAMDLILNSLPLVSDAETSLTICIASGWPHPEPTIGRD 60
QY 61 FEALMNHQDPLEVTQVTPEWAKVVKPEKASKINGAYFCGRVPGEAIPRTMKMPQ 120
DB 61 FEALMNHQDPLEVTQVTPEWAKVVKPEKASKINGAYFCGRVPGEAIPRTMKMPQ 120
QY 121 QASFLPATLMTVDKGNVNSIFSKVLIKEEDAVIYKNGSFIHSVPHEVEDILEVHLPH 180
DB 121 QASFLPATLMTVDKGNVNSIFSKVLIKEEDAVIYKNGSFIHSVPHEVEDILEVHLPH 180
QY 181 AQPDAGVYSARYIGGNLFSAFTKILVRKCEAKWKPCPNHCTACMNNGVCHEDIGEC 240

```


QY 578 CKVSKALPAIEKTIKAKGQPPQVYTLPPSPPEMTKQVSIITLVKGFPSDIAVE 637
 DB 204 CKVSKALPAIEKTIKAKGQPPQVYTLPPSPPEMTKQVSIITLVKGFPSDIAVE 263
 QY 648 WESNGVPENNYKTIIPVLSGSEFFLYSKLIVDKSRWQGNVFSVMEALHNHYTQKS 697
 DB 264 WESNGVPENNYKTIIPVLSGSEFFLYSKLIVDKSRWQGNVFSVMEALHNHYTQKS 323
 QY 698 LSLSPCK 704
 DB 324 LSLSPCK 330

RESULT 8
 S41866
 Iq gamma-3 chain C region - synthetic
 C:Species: synthetic
 A:Note: Homo sapiens (man) gene engineered and expressed in *Escherichia coli*
 C:Date: 06-Jan-1995 #sequence_revision 17-Mar-1997 #text_change 19-May-2000
 C:Accession: S41866
 R:Filipula, D.
 submitted to the EMBL Data Library, February 1993
 A:Description: Screening method for protein-protein interactions of cloned gene products.
 A:Reference number: S41866
 A:Accession: S41866
 A:Molecule type: mRNA
 A:Residues: 1-255 +Flg
 A:Cross references: EMBL:X70421; NID:33068; PDB:CAA9866.1; PDB:33069
 C:Keywords: immunoglobulin
 E:1-255/Region: *Escherichia coli* outer membrane protein A precursor
 F:24-255/Region: human Iq gamma 1 chain C region

Query Match 32.3% Score 1249. ID: 4; Length 255;
 Best Local Similarity 98.7% Pred. No. 3, 8e-63;
 Matches 229; Conservative 2; Mismatches 1; Indels 0; Gaps 0.

QY 473 EHKSCDKHTCHCPDPAPELGGPSVFLPPPKDILMISRTPEVTVVIVSHEDP 542
 DB 24 EHKSCDKHTCHCPDPAPELGGPSVFLPPPKDILMISRTPEVTVVIVSHEDP 83
 QY 533 NNYVDGVEVHNAKTPREEQYNSTYRVVSVITVLHQWLNGKEYKCKVSKALPAIEKT 592
 DB 84 NNYVDGVEVHNAKTPREEQYNSTYRVVSVITVLHQWLNGKEYKCKVSKALPAIEKT 143
 QY 593 ISKAKGQPPQVYTLPPSPPEMTKQVSIITLVKGFPSDIAVEWESNGQPNNYKTIIP 652
 DB 144 ISKAKGQPPQVYTLPPSPPEMTKQVSIITLVKGFPSDIAVEWESNGQPNNYKTIIP 203
 QY 653 PVLSDGSEFFLYSKLIVDKSRWQGNVFSVMEALHNHYTQKSLSLSPCK 704
 DB 204 PVLSDGSEFFLYSKLIVDKSRWQGNVFSVMEALHNHYTQKSLSLSPCK 255

RESULT 9
 P10207
 Iq gamma chain C region - chimpanzee
 C:Species: Pan troglodytes (chimpanzee)
 C:Date: 24-Nov-1991 #sequence_revision 23-Nov-1991 #text_change 16-Jul-1994
 C:Accession: P10207
 R:Erlich, F.H.; Moustafa, Z.A.; Oestergaard, L.
 Mol. Immunol. 28, 319-322, 1991
 A:Title: Nucleotide sequence of chimpanzee Fc and hinge regions.
 A:Reference number: P10207; MUID:91287716
 A:Accession: P10207
 A:Molecule type: mRNA
 A:Residues: 1-234 +EHR
 C:Superfamily: immunoglobulin C region; immunoglobulin homology
 C:Keywords: immunoglobulin
 F:48-117/Domain: immunoglobulin homology <IMM>

Query Match 31.1% Score 1204. ID: 2; Length 244;
 Best Local Similarity 97.8% Pred. No. 1, 2e-60;
 Matches 220; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 473 EHKSCDKHTCHCPDPAPELGGPSVFLPPPKDILMISRTPEVTVVIVSHEDP 542
 DB 10 EHKSCDKHTCHCPDPAPELGGPSVFLPPPKDILMISRTPEVTVVIVSHEDP 69
 QY 533 NNYVDGVEVHNAKTPREEQYNSTYRVVSVITVLHQWLNGKEYKCKVSKALPAIEKT 592
 DB 70 NNYVDGVEVHNAKTPREEQYNSTYRVVSVITVLHQWLNGKEYKCKVSKALPAIEKT 129
 QY 593 ISKAKGQPPQVYTLPPSPPEMTKQVSIITLVKGFPSDIAVEWESNGQPNNYKTIIP 652
 DB 130 ISKAKGQPPQVYTLPPSPPEMTKQVSIITLVKGFPSDIAVEWESNGQPNNYKTIIP 189
 QY 653 PVLSDGSEFFLYSKLIVDKSRWQGNVFSVMEALHNHYTQKS 697
 DB 160 PVLSDGSEFFLYSKLIVDKSRWQGNVFSVMEALHNHYTQKS 244

RESULT 10
 A23511
 Iq gamma-3 chain C region (allotype G3m(b)) - human
 C:Species: Homo sapiens (man)
 C:Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 23-Jul-1999
 C:Accession: A23511
 R:Huck, S.; Fort, P.; Crawford, D.H.; LeFranc, M.P.; LeFranc, G.
 Nucleic Acids Res. 14, 1779-1789, 1986
 A:Title: Sequence of a human immunoglobulin gamma 3 heavy chain constant region gene.
 A:Reference number: A23511; MUID:86148507
 A:Accession: A23511
 A:Molecule type: DNA
 A:Residues: 1-377 +HUC
 A:Cross references: GH:X03604; GH:M12958; NID:33070; PDB:CAA27268.1; PDB:4577056
 C:Keywords: immunoglobulin homology <IMM>

Query Match 30.6% Score 1183. ID: 2; Length 477;
 Best Local Similarity 92.4% Pred. No. 2, 9e-59;
 Matches 218; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

QY 469 PVFPEPKSVDKTHITPPDPAPELGGPSVFLPPPKDILMISRTPEVTVVIVSHEDP 628
 DB 142 PVFPEPKSVDKTHITPPDPAPELGGPSVFLPPPKDILMISRTPEVTVVIVSHEDP 201
 QY 529 EVKFNWYVDGVEVHNAKTPREEQYNSTYRVVSVITVLHQWLNGKEYKCKVSKALPAIEKT 588
 DB 202 EVKFNWYVDGVEVHNAKTPREEQYNSTYRVVSVITVLHQWLNGKEYKCKVSKALPAIEKT 261
 QY 589 HKTTISKAKGQPPQVYTLPPSPPEMTKQVSIITLVKGFPSDIAVEWESNGQPNNY 648
 DB 262 HKTTISKAKGQPPQVYTLPPSPPEMTKQVSIITLVKGFPSDIAVEWESNGQPNNY 721
 QY 649 KTTPEVLSGSEFFLYSKLIVDKSRWQGNVFSVMEALHNHYTQKSLSLSPCK 704
 DB 322 KTTPEVLSGSEFFLYSKLIVDKSRWQGNVFSVMEALHNHYTQKSLSLSPCK 477

RESULT 11
 A60764
 Iq gamma-3 chain C region, form IAI - human
 C:Species: Homo sapiens (man)
 C:Date: 14-May-1993 #sequence_revision 14-May-1993 #text_change 16-Jul-1999
 C:Accession: A60764
 R:Huck, S.; LeFranc, G.; LeFranc, M.P.

Immunogenetics 30, 250-257, 1989
 A:Title: A human immunoglobulin IGHG3 allele (emb6, b1, c3, c5, u) with an IGHG4 convert
 A:Reference number: A60764; MUID:90007613
 A:Accession: A60764
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-377 <H2C>
 C:Superfamily: immunoglobulin c region; immunoglobulin homology
 C:Keywords: immunoglobulin
 F:20-85/Domain, immunoglobulin homology <IMM>

Query Match 30.9%; Score 1191; DB 2; Length 377;
 Best Local Similarity 92.4%; Pred. No. 3 75-59;
 Matches 218; Conservative 6; Mismatches 12; Indels 0; Gaps 0.

Qy 469 PYFPEKSKNTKTPCPAPPELLGGPSVFLPPKPKDILMSPTPEVTCVVVDVSHEDP 528
 Db 142 PRCPPEKSCDTPPCPCAPPELLGGPSVFLPPKPKDILMSPTPEVTCVVVDVSHEDP 201
 Qy 529 EVKFNWVGVGVHNAKTPKPKQYNSTFRVSVLVTHQDWLNCKEYKCKVSNKALPAP 588
 Db 202 EVQPKWVGVGVHNAKTPKPKQYNSTFRVSVLVTHQDWLNCKEYKCKVSNKALPAP 261
 Qy 589 IEKTSKAKGQPREPQVYTLPPSRPEEMTKNOVSLTCLVKGFYPSDIAVWESNGQPENNY 648
 Db 262 IEKTSKAKGQPREPQVYTLPPSRPEEMTKNOVSLTCLVKGFYPSDIAVWESNGQPENNY 321
 Qy 649 KITPVLDSGSFELYSLKLVDSKRWQGNVFSVSNVHEALNHVYTKSLISLSPG 704
 Db 322 NITPVLDSGSFELYSLKLVDSKRWQGNVFSVSNVHEALNHVYTKSLISLSPG 377

RESULT 12

G3HUI

Ig gamma-3 heavy chain disease proteins - human
 C:Species: Homo sapiens (man)
 C:Date: 31-Dec-1979 #sequence_revision 21 Jun 1983 #text_change 16-Jul-1999
 C:Accession: A90442; A92219; A93915; A92149
 R:Frangione, B.; Rosenwasser, E.; Prelli, F.; Franklin, E.C.
 Biochemistry 19, 4304-4308, 1980
 A:Title: Primary structure of human gamma3 immunoglobulin deletion mutant: gamma3 heavy-
 A:Reference number: A90442; MUID:81021548
 A:Contents: heavy chain disease protein Wis
 A:Accession: A90442

A:Molecule type: protein
 A:Residues: 1-289 <PFA>
 A:Note: The molecule is a dimer linked by 12 disulfide bonds; it has an extra interchain
 A:Note: this protein lacks most of the V region and all of the C_H1 region. Residue 12 of
 A:Note: the sequence of residues 42-76 was taken from the reference that follows
 R:Michaels, T.E.; Frangione, B.; Franklin, E.C.
 J. Biol. Chem. 252, 883-889, 1977

A:Title: Primary structure of the 'hinge' region of human IgG3. Probably quadruplication
 A:Reference number: A92219; MUID:7711856;
 A:Contents: normal gamma-3 chains, sequence corresponding to residues 12-97 of protein W

A:Accession: A92219
 A:Molecule type: protein
 A:Residues: 12-97 <MC>

A:Note: The hinge region in gamma-3 chains is about four times as long as in other gamma
 A:Note: segments (12-28)

R:Wolfeinstein-Tudel, G.; Frangione, B.; Prelli, F.; Franklin, E.C.
 Biochem Biophys Res Commun 71, 907-914, 1976

A:Title: The amino acid sequence of "heavy chain disease" protein 20C. Structure of the
 A:Reference number: A90198; MUID:77021516

A:Contents: heavy chain disease protein 20C, partial sequence corresponding to residues
 A:Accession: A90198

A:Molecule type: protein
 A:Residues: 59-125, 187, 128-226, 228-289 <WOL>

A:Note: This protein lacks most of the V region, all of the C_H1 region, and part of the
 R:Alexander, A.; Steinmetz, M.; Barillet, D.; Frangione, B.; Franklin, E.C.; Hood, L.

Proc Natl Acad Sci U.S.A. 79, 3260-3264, 1982
 A:Title: gamma heavy chain disease in man: cDNA sequence supports partial gene deletion

A:Reference number: A93915; MUID:82247835
 A:Contents: heavy chain disease protein Omn
 A:Accession: A93915
 A:Molecule type: mRNA
 A:Residues: 12 70,72 114,116 125, 127,133, 135,138, 139,140-154, 155-156-
 A:Note: a carboxyl terminal Lys is removed posttranslationally
 C:Superfamily: immunoglobulin gamma chain subclass
 C:Comment: This sequence may represent an allelic form of another gamma chain subclass
 A:Note: The heavy chain disease protein Wis is shown.

C:Keywords: immunoglobulin

F:16,140/Binding site: carbohydrate (Asn) (covalent) #status experimental

A:Gene: GDB:IGHG3

A:Cross-references: GDB:119339; OMIM:147120

A:Map position: 14q32.33-14q32.33

C:Superfamily: immunoglobulin C region, immunoglobulin homology/
 C:Keywords: duplication; glycoprotein, immunoglobulin, pyridinamic acid

F:203-270/Domain, immunoglobulin homology <IMM>

F:1/Modified site: pyridone carboxylic acid (Gln) #status experimental

F:6,140/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 29.9%; Score 1158; DB 1; Length 289;
 Best Local Similarity 90.2%; Pred. No. 5,4e-58;
 Matches 212; Conservative 11; Mismatches 12; Indels 0; Gaps 0;

Qy 469 PYFPEKSKNTKTPCPAPPELLGGPSVFLPPKPKDILMSPTPEVTCVVVDVSHEDP 528
 Db 55 PRPEKSCDTPPCPCAPPELLGGPSVFLPPKPKDILMSPTPEVTCVVVDVSHEDP 114
 Qy 529 EVKFNWVGVGVHNAKTPKPKQYNSTFRVSVLVTHQDWLNCKEYKCKVSNKALPAP 588
 Db 115 EVQPKWVGVGVHNAKTPKPKQYNSTFRVSVLVTHQDWLNCKEYKCKVSNKALPAP 174
 Qy 589 IEKTSKAKGQPREPQVYTLPPSRPEEMTKNOVSLTCLVKGFYPSDIAVWESNGQPENNY 648
 Db 175 IEKTSKAKGQPREPQVYTLPPSRPEEMTKNOVSLTCLVKGFYPSDIAVWESNGQPENNY 234
 Qy 649 KITPVLDSGSFELYSLKLVDSKRWQGNVFSVSNVHEALNHVYTKSLISLSPG 703
 Db 234 NITPVLDSGSFELYSLKLVDSKRWQGNVFSVSNVHEALNHVYTKSLISLSPG 289

RESULT 13

G2HU

Ig gamma-2 chain C region human

C:Species: Homo sapiens (man)

C:Date: 30-Apr-1981 #sequence_revision 13 Jun 1983 #text_change 21-Jul-2000

C:Accession: A93906; A92809; A90752; A93132; A02148

R:Ellison, J., Hood, L.

Proc Natl Acad Sci U.S.A. 79, 1984 1988, 1982

A:Title: Linkage and sequence homology of two human immunoglobulin gamma heavy chain

A:Reference number: A93906; MUID:82197621

A:Accession: A93906

A:Molecule type: DNA

A:Residues: 1-325 <ALL>

A:Cross-references: GDB:60654; GDB:108233; NID-94759; FIUN:CA258436.1; FID:4566056

A:Note: Lys-326 is probably removed posttranslationally

R:Wang, A.C.; Tuzi, E.; Fudenberg, H.H.

J. Immunol. 125, 1048-1054, 1980

A:Title: The primary structure of a human IgG2 heavy chain: genetic, evolutionary, an

A:Reference number: A92809; MUID:81007873

A:Contents: myeloma protein T11

A:Accession: A92809

A:Molecule type: protein

A:Residues: 1-19, 20, 21, 27, 29, 31, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101, 102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127, 128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139, 140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152, 153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167, 168, 169, 170, 171, 172, 173, 174, 175, 176, 177, 178, 179, 180, 181, 182, 183, 184, 185, 186, 187, 188, 189, 190, 191, 192, 193, 194, 195, 196, 197, 198, 199, 200, 201, 202, 203, 204, 205, 206, 207, 208, 209, 210, 211, 212, 213, 214, 215, 216, 217, 218, 219, 220, 221, 222, 223, 224, 225, 226, 227, 228, 229, 230, 231, 232, 233, 234, 235, 236, 237, 238, 239, 240, 241, 242, 243, 244, 245, 246, 247, 248, 249, 250, 251, 252, 253, 254, 255, 256, 257, 258, 259, 260, 261, 262, 263, 264, 265, 266, 267, 268, 269, 270, 271, 272, 273, 274, 275, 276, 277, 278, 279, 280, 281, 282, 283, 284, 285, 286, 287, 288, 289, 290, 291, 292, 293, 294, 295, 296, 297, 298, 299, 300, 301, 302, 303, 304, 305, 306, 307, 308, 309, 310, 311, 312, 313, 314, 315, 316, 317, 318, 319, 320, 321, 322, 323, 324, 325, 326, 327, 328, 329, 330, 331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 344, 345, 346, 347, 348, 349, 350, 351, 352, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375, 376, 377, 378, 379, 380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, 400, 401, 402, 403, 404, 405, 406, 407, 408, 409, 410, 411, 412, 413, 414, 415, 416, 417, 418, 419, 420, 421, 422, 423, 424, 425, 426, 427, 428, 429, 430, 431, 432, 433, 434, 435, 436, 437, 438, 439, 440, 441, 442, 443, 444, 445, 446, 447, 448, 449, 450, 451, 452, 453, 454, 455, 456, 457, 458, 459, 460, 461, 462, 463, 464, 465, 466, 467, 468, 469, 470, 471, 472, 473, 474, 475, 476, 477, 478, 479, 480, 481, 482, 483, 484, 485, 486, 487, 488, 489, 490, 491, 492, 493, 494, 495, 496, 497, 498, 499, 500, 501, 502, 503, 504, 505, 506, 507, 508, 509, 510, 511, 512, 513, 514, 515, 516, 517, 518, 519, 520, 521, 522, 523, 524, 525, 526, 527, 528, 529, 530, 531, 532, 533, 534, 535, 536, 537, 538, 539, 540, 541, 542, 543, 544, 545, 546, 547, 548, 549, 550, 551, 552, 553, 554, 555, 556, 557, 558, 559, 560, 561, 562, 563, 564, 565, 566, 567, 568, 569, 570, 571, 572, 573, 574, 575, 576, 577, 578, 579, 580, 581, 582, 583, 584, 585, 586, 587, 588, 589, 590, 591, 592, 593, 594, 595, 596, 597, 598, 599, 600, 601, 602, 603, 604, 605, 606, 607, 608, 609, 610, 611, 612, 613, 614, 615, 616, 617, 618, 619, 620, 621, 622, 623, 624, 625, 626, 627, 628, 629, 630, 631, 632, 633, 634, 635, 636, 637, 638, 639, 640, 641, 642, 643, 644, 645, 646, 647, 648, 649, 650, 651, 652, 653, 654, 655, 656, 657, 658, 659, 660, 661, 662, 663, 664, 665, 666, 667, 668, 669, 670, 671, 672, 673, 674, 675, 676, 677, 678, 679, 680, 681, 682, 683, 684, 685, 686, 687, 688, 689, 690, 691, 692, 693, 694, 695, 696, 697, 698, 699, 700, 701, 702, 703, 704, 705, 706, 707, 708, 709, 710, 711, 712, 713, 714, 715, 716, 717, 718, 719, 720, 721, 722, 723, 724, 725, 726, 727, 728, 729, 730, 731, 732, 733, 734, 735, 736, 737, 738, 739, 740, 741, 742, 743, 744, 745, 746, 747, 748, 749, 750, 751, 752, 753, 754, 755, 756, 757, 758, 759, 760, 761, 762, 763, 764, 765, 766, 767, 768, 769, 770, 771, 772, 773, 774, 775, 776, 777, 778, 779, 780, 781, 782, 783, 784, 785, 786, 787, 788, 789, 790, 791, 792, 793, 794, 795, 796, 797, 798, 799, 800, 801, 802, 803, 804, 805, 806, 807, 808, 809, 810, 811, 812, 813, 814, 815, 816, 817, 818, 819, 820, 821, 822, 823, 824, 825, 826, 827, 828, 829, 830, 831, 832, 833, 834, 835, 836, 837, 838, 839, 840, 841, 842, 843, 844, 845, 846, 847, 848, 849, 850, 851, 852, 853, 854, 855, 856, 857, 858, 859, 860, 861, 862, 863, 864, 865, 866, 867, 868, 869, 870, 871, 872, 873, 874, 875, 876, 877, 878, 879, 880, 881, 882, 883, 884, 885, 886, 887, 888, 889, 890, 891, 892, 893, 894, 895, 896, 897, 898, 899, 900, 901, 902, 903, 904, 905, 906, 907, 908, 909, 910, 911, 912, 913, 914, 915, 916, 917, 918, 919, 920, 921, 922, 923, 924, 925, 926, 927, 928, 929, 930, 931, 932, 933, 934, 935, 936, 937, 938, 939, 940, 941, 942, 943, 944, 945, 946, 947, 948, 949, 950, 951, 952, 953, 954, 955, 956, 957, 958, 959, 960, 961, 962, 963, 964, 965, 966, 967, 968, 969, 970, 971, 972, 973, 974, 975, 976, 977, 978, 979, 980, 981, 982, 983, 984, 985, 986, 987, 988, 989, 990, 991, 992, 993, 994, 995, 996, 997, 998, 999, 1000

R:Hotmann, T.; Parr, D.M.
Mol. Immunol. 16, 923-925, 1979
A:Title: A note on the amino acid sequence of residues 381-391 of human immunoglobulin G
A:Reference number: A93142; MUID:80114419
A:Contents: Zie
A:Accession: A93132
A:Molecule type: protein
A:Residues: 238-275, H3P
R:Hotmann, T.; Parr, D.M.
submitted to the Atlas, March 1980
A:Reference number: A94591
A:Contents: annotation; 216 revisions to residues 26, 69, 68, and 264-268
A:Note: the revised sequence differs from that shown in having 60 Ala and in the amidated
residues
R:Milstein, C.; Franquone, B.
Biochem. J. 121, 217-225, 1971
A:Title: Disulphide bridges of the heavy chain of human immunoglobulin G2.
A:Reference number: A90280; MUID:70203360
A:Contents: annotation; myeloma protein Sa, disulfide bonds
R:Frangione, B.; Milstein, C.; Pink, J.R.L.
Nature 221, 145-148, 1969
A:Title: Structural studies of immunoglobulin G
A:Reference number: A93157; MUID:69064124
A:Contents: annotation; Sa, disulfide bonds
C:Genetics:
A:Gene: GMB, IGHC2
A:Cross-references: GDB:119348; OMIM:147110
A:Map position: 14332.33 14332.73
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (Lg) and heavy (Hg) chains. In some cases, such as IgA and IgM, the subunits associate into higher order structures.
C:Superfamily: immunoglobulin C region, immunoglobulin homology
C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
F:20-85/Domain: immunoglobulin homology <IM2>
F:134-203/Domain: immunoglobulin homology <IM3>
F:234-306/Domain: immunoglobulin homology <IM3>
F:14/Disulfide bonds: interchain (to light chain) #status experimental
F:27-83,141-201,247-305/Disulfide bonds: #status experimental
F:102,103,106,109/Disulfide bonds: interchain (to heavy chain) #status experimental
F:176/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 29.8% Score 1154 DB 1: Length 326
Best Local Similarity 82.4% Pred. No. 1e-57
Matches 215, Conservative 13, Mismatches 15, Gaps 2.

QY 461 AVINISSEYFGEPSKUIH - - - - - TCCPCAPPELLGSPVFLPPPK 503
DB 67 SVTVVSSSEFGTQYTCNWDKPSIKVKDKVERKKQCECHPCAPP-VAGSPVFLPPPK 125
QY 504 PKDTLMSIRPEPTCVVWVSHEDPEVFNWYVGVVEVHNKIKPREQYNSTYRVVSVL 563
DB 126 PKDTLMSIRPEPTCVVWVSHEDPEVFNWYVGVVEVHNKIKPREQYNSTYRVVSVL 185
QY 564 IVHQWLNGKPKYKVKVNNALPAPTEKTSKAKGDPPEPVYTLPPSPPEMKNQVSLI 623
DB 186 IVVHQWLNGKPKYKVKVNNKGLPAPTEKTSKKGQPEPVYTLPPSPPEMKNQVSLT 245
QY 624 CLVKGYFSDIAVENESNCPENNYKTTPTFVLDSGSEFLYSLKTLVDKSRWQNVLS 683
DB 246 CLVKGYFSDIAVENESNCPENNYKTTPTFVLDSGSEFLYSLKTLVDKSRWQNVLS 305
QY 684 VMHEALHNHYTKSLSPGK 704
DB 306 VMHEALHNHYTKSLSPGK 326

RESULT 14
GHRH
Ig gamma 4 chain C region - human
C:Species: Homo sapiens (man)
C:Date: 02-Apr-1982 #sequence_revision 92 Apr-1982 #text_change 16-Jul-1999
C:Accession: A90933; A90249; A92150
R:Ellison, J.; Huxbaum, J.; Hood, L.

DNA 1, 11-18, 1981
A:Title: Nucleotide sequence of a human immunoglobulin C gamma4 gene.
A:Reference number: A90933; MUID:83157104
A:Accession: A90933
A:Molecule type: DNA
A:Residues: 1-327 <ELL>
A:Note: the sequence was determined from the germline gene.
R:Pink, J.R.L.; Buttery, S.H.; De Vries, G.M.; Milstein, C.
Biochem. J. 117, 33-47, 1970
A:Title: Human immunoglobulin subclasses. Partial amino acid sequence of the constant region of the heavy chain of human immunoglobulin G2.
A:Reference number: A90249; MUID:70207560
A:Accession: A90249
A:Molecule type: protein
A:Residues: 1-30; 81-326 <PN>
C:Genetics:
A:Gene: GDB, IGHC4
A:Cross-references: GDB:119340; OMIM:147110
A:Map position: 14332.33 14332.33
A:Introns: 99/1; 111/1; 221/1
C:Complex: An immunoglobulin heterotetramer subunit consists of two identical light (Lg) and heavy (Hg) chains. In some cases, such as IgA and IgM, the subunits associate into higher order structures.
C:Superfamily: immunoglobulin C region, immunoglobulin homology
C:Keywords: duplication; glycoprotein; heterotetramer; immunoglobulin
F:20-85/Domain: immunoglobulin homology <IM2>
F:99-110/Region: hinge
F:134-203/Domain: immunoglobulin homology <IM2>
F:240-307/Domain: immunoglobulin homology <IM3>
F:14/Disulfide bonds: interchain (to light chain) #status experimental
F:27-83,141-201,247-305/Disulfide bonds: #status predicted
F:106,109/Disulfide bonds: interchain (to heavy chain) #status experimental
F:177/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 29.6% Score 1144 DB 1: Length 327
Best Local Similarity 89.9% Pred. No. 3.8e-57
Matches 213, Conservative 7, Mismatches 9, Gaps 1;

QY 468 HYPFEPKSKUIHCHPCAPPELLGSPVFLPPPKKUIHMSRPEVTVVWVSHED 527
DB 99 FSKYGP - - - - - CPSCPAPEFLGSPVFLPPPKKUIHMSRPEVTVVWVSHED 150
QY 528 PVKFNWYVGVVEVHNKIKPREQYNSTYRVVSVLIVHQWLNGKPKYKVKVNNALPA 587
DB 151 PVQFNWYVGVVEVHNKIKPREQYNSTYRVVSVLIVHQWLNGKPKYKVKVNNALPA 210
QY 588 PTEKTSKAKGDPPEPVYTLPPSPPEMKNQVSLTCLVKGYFSDIAVENESNCPEN 647
DB 211 STEKTSKAKGDPPEPVYTLPPSPPEMKNQVSLTCLVKGYFSDIAVENESNCPEN 270
QY 648 YKTTPTVLDSDGSEFLYSLKTLVDKSRWQNVFSCVMHEALHNHYTKSLSPGK 704
DB 271 YKTTPTVLDSDGSEFLYSLKTLVDKSRWQNVFSCVMHEALHNHYTKSLSPGK 327

RESULT 15
GHRH
Ig gamma chain C region - rabbit
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 24-Apr-1984 #sequence_revision 15-Nov-1984 #text_change 16-Jul-1999
C:Accession: A91749; A90240; A90245; A94416; A92161
R:Bernstein, K.E.; Alexander, C.B.; Mage, R.G.
Immunogenetics 18, 387-397, 1983
A:Title: Nucleotide sequence of a rabbit IgG heavy chain from the recombinant F1 hap
A:Reference number: A91749; MUID:84030930
A:Accession: A91749
A:Molecule type: mRNA
A:Residues: 1-323 <HER>
A:Note: this sequence has the 412 alleotypic marker, 104-107, and the 614 marker, 145-146.
R:Pratt, D.M.; Mole, L.E.
Biochem. J. 151, 337-349, 1975
A:Title: Sequence studies on the constant region of the Fd sections of rabbit immunoglobulin G2.
A:Reference number: A90290; MUID:76135469
A:Accession: A90290

A:Molecule type: protein
 A:Residues: 1-47, 'E', 49-71, 'FV', 72-128 <FEA>
 R:Martens, C.L.; Moore, K.W.; Steinmetz, M.; Hood, L.; Knight, K.L.
 Proc. Natl. Acad. Sci. U.S.A. 79: 6018-6022, 1982
 A:Title: Heavy chain genes of rabbit IgG: Isolation of a cDNA encoding gamma heavy chain
 A:Reference number: A93928; MUID:83399917
 A:Accession: A93928
 A:Molecule type: mRNA
 A:Residues: 48-103, 'M', 105-143, 'P', 145-184, 'A', 186, 'E', 198-256 <MAR>
 A:Cross-references: CR M16426; NID-g165111; P1IN-AAA11289.1; P1D-g165112
 A:Note: this sequence has the d11 allotypic marker, 104-Met, and the e15 allotypic marker
 R:Euchner, R.G.; Jackson, S.A.; Mole, L.E.; Porter, P.P.
 Biochem. J. 116, 249-259, 1970
 A:Title: Sequence studies of the Fd section of the heavy chain of rabbit immunoglobulin
 A:Reference number: A90245; MUID:70110015
 A:Accession: A90245
 A:Molecule type: protein
 A:Residues: 132-143, 'E', 145-161 <FRD>
 R:Hall, R.L.; Lebovitz, H.E.; Fellows Jr., R.E.; Delaney, R.
 In Gamma Globulins, N-Ed Sym. A, Elland-1, 1, ed., pp. 109-127, Almqvist and Wiksell,
 A:Reference number: A94416
 A:Accession: A94416
 A:Molecule type: protein
 A:Residues: 129-131, 155-172, 'D', 174-184, 'A', 186, 'E', 188-200, 'D', 202-212, 'E', 219-232, 'Q',
 A:Note: this has the e15 allotypic marker, 185-Ala
 C:Complex. An immunoglobulin heterotetramer subunit consists of two identical light (kap
 hair disulfide bonds. In some cases, such as IgA and IgM, the subunits associate into la
 C:Superfamily. Immunoglobulin C region; immunoglobulin homology
 C:Keywords. duplication, glycoprotein, heterotetramer; immunoglobulin
 F:26-82/Domains: immunoglobulin homology <IM1>
 F:130-169/Domains: immunoglobulin homology <IM2>
 F:235-303/Domains: immunoglobulin homology <IM3>
 F:173/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 24.1%; Score 932; DB 1; Length 323;
 Best Local Similarity 59.5%; Pred. No. 2.8e-45;
 Matches 173; Conservative 40; Mismatches 48; Indels 30; Gaps 3;
 QY 444 LPKPLNA-----PNI DTGHNFAVINISSEPFGEKPKSCDKTH----- 481
 DB 33 LPEPVTITWNSGLTNGVTFPPSPQSSGLYSLSSVSVTSQQVTCNVNHPATNTKVD 92
 QY 482 -----TC--PPCPAPELLGGSPVELFPKPKDITLMSIRTPETCVVVDVSHEDPEVKEN 533
 DB 93 KTVAPSTCSKPTCTPPPELGGSPVFIIPPAPKDTLMSIRTPETCVVVDVSDQDDPEVQT 152
 QY 534 WYDGVVEVHNAKTPREQYNSTYRVVSVLTVLHODWLNKGYKCKVSKNKPAPIEKTI 593
 DB 153 WYINNEQVRTAPPPIPPQFNSTIPVVSITLPI THODWLNKGYKCKVSKNKPAPIEKTI 212
 QY 594 SKAGQPREPOVYIIPPSREMTKNQVSLICLVKGYPSDIAVWFSNGOPENNYKTPPP 653
 DB 213 SKAFGPLEKPVYIMPPPEELSSPSVSLDMINGEYPSDLSVWEKFNKAENYKTTFA 272
 QY 654 VLSDSGSFELYSLTVKDSKQGNVPSGSMHEALHNHYTKQSLSPGK 704
 DB 273 VLSDSGSYFLYNKLSVPTSHQKQDVPTCSVMHEALHNHYTKQSLSPGK 323

Search completed: September 23, 2002, 22:11:26
 Job time: 162 sec

100

101

102

103

104

105

106

107

108

109

110

111

112

113

114

115

116

117

118

119

120

121

122

123

124

125

126

127

128

129

130

131

132

133

134

135

136

137

138

139

140

141

142

143

144

145

146

147

148

149

150

151

152

153

154

155

156

157

158

159

160

161

162

163

164

165

166

167

168

169

170

171

172

173

174

175

176

177

178

179

180

181

182

183

184

185

186

187

188

189

190

191

192

193

194

195

196

197

198

199

200

201

202

203

204

205

206

207

208

209

210

211

212

213

214

215

216

217

218

219

220

221

222

223

224

225

226

227

228

229

230

231

232

233

234

235

236

237

238

239

240

241

242

243

244

245

246

247

248

249

250

251

252

253

254

255

256

257

258

259

260

261

262

263

264

265

266

267

268

269

270

271

272

273

274

275

276

277

278

279

280

281

282

283

284

285

286

287

288

289

290

291

292

293

294

295

296

297

298

299

300

301

302

303

304

305

306

307

308

309

310

311

312

313

314

315

316

317

318

319

320

321

322

323

324

325

326

327

328

329

330

331

332

333

334

335

336

337

338

339

340

341

342

343

344

345

346

347

348

349

350

351

352

353

354

355

356

357

358

359

360

361

362

363

364

365

366

367

368

369

370

371

372

373

374

375

376

377

378

379

380

381

382

383

384

385

386

387

388

389

390

391

392

393

394

395

396

397

398

399

400

401

402

403

404

405

406

407

408

409

410

411

412

413

414

415

416

417

418

419

420

421

422

423

424

425

426

427

428

429

430

431

432

433

434

435

436

437

438

439

440

441

442

443

444

445

446

447

448

449

450

451

452

453

454

455

456

457

458

459

460

461

462

463

464

465

466

467

468

469

470

471

472

473

474

475

476

477

478

479

480

481

482

483

484

485

486

487

488

489

490

491

492

493

494

495

496

497

498

499

500

501

502

503

504

505

506

507

508

509

510

511

512

513

514

515

516

517

518

519

520

521

522

523

524

525

526

527

528

529

530

531

532

533

534

535

536

537

538

539

540

541

542

543

544

545

546

547

548

549

550

551

552

553

554

555

556

557

558

559

560

561

562

563

564

565

566

567

568

569

570

571

572

573

574

575

576

577

578

579

580

581

582

583

584

585

586

587

588

589

590

591

592

593

594

595

596

597

598

599

600

601

602

603

604

605

606

607

608

609

610

611

612

613

614

615

616

617

618

619

620

621

622

623

624

625

626

627

628

629

630

631

632

633

634

635

636

637

638

639

640

641

642

643

644

645

646

647

648

649

650

651

652

653

654

655

656

657

658

659

660

661

662

663

664

665

666

667

668

669

670

671

672

673

674

675

676

677

678

679

680

681

682

683

684

685

686

687

688

689

690

691

692

693

694

695

696

697

698

699

700

701

702

703

704

705

706

707

708

709

710

711

712

713

714

715

716

717

718

719

720

721

722

723

724

725

726

727

728

729

730

731

732

733

734

735

736

737

738

739

740

741

742

743

744

745

746

747

748

749

750

751

752

753

754

755

756

757

758

759

760

761

762

763

764

765

766

767

768

769

770

771

772

773

774

775

776

777

778

779

780

781

782

783

784

785

786

787

788

789

790

791

792

793

794

795

796

797

798

799

800

801

802

803

804

805

806

807

808

809

810

811

812

813

814

815

816

817

818

819

820

821

822

823

824

825

826

827

828

829

830

831

832

833

834

835

836

837

838

839

840

841

842

843

844

845

846

847

848

849

850

851

852

853

854

855

856

857

858

859

860

861

862

863

864

865

866

867

868

869

870

871

872

873

874

875

876

877

878

879

880

881

882

883

884

885

886

887

888

889

890

891

892

893

894

895

896

897

898

899

900

901

902

903

904

905

906

907

908

909

910

911

912

913

914

915

916

917

918

919

920

921

922

923

924

925

926

927

928

929

930

931

932

933

934

935

936

937

938

939

940

941

942

943

944

945

946

947

948

949

950

951

952

953

954

955

956

957

958

959

960

961

962

963

964

965

966

967

968

969

970

971

972

973

974

975

976

977

978

979

980

981

982

983

984

985

986

987

988

989

990

991

992

993

994

995

996

997

998

999

1000

CC MAMMALIAN ENDOTHELIAL CELL LINEAGE MARKER. PROBABLY REGULATES
 CC ENDOTHELIAL CELL PROLIFERATION, DIFFERENTIATION AND GUIDES THE
 CC PROPER PATTERNING OF ENDOTHELIAL CELLS DURING BLOOD VESSEL
 CC FORMATION.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein
 CC -1- TISSUE SPECIFICITY: SPECIFICALLY EXPRESSED IN DEVELOPING VASCULAR
 CC ENDOTHELIAL CELLS.
 CC -1- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
 CC DOMAIN.
 CC -1- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 3 FIBRONECTIN TYPE III-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL, collaboration
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X71424; CAA50555.1; -
 CC PIR: S32691; S32691.
 CC HSP: P11362; LFGR.
 CC InterPro: IPR000561; EGF like.
 CC InterPro: IPR000719; Euk_pkinase.
 CC InterPro: IPR003961; FN_III.
 CC InterPro: IPR002049; Laminin_EGF.
 CC InterPro: IPR001245; Tyr_pkinase.
 CC Pfam: PF00008; EGF; 2.
 CC Pfam: PF00041; fn3; 3.
 CC Pfam: PF00069; pkinase; 1.
 CC PRINTS: PR00109; TYRKINASE.
 CC SMART: SM00180; EGF_Lam; 1.
 CC SMART: SM00001; EGF_Like; 1.
 CC SMART: SM00060; FN3; 3.
 CC SMART: SM00219; TYRK; 1.
 CC PROSITE: PS00107; PROTEIN_KINASE_1.
 CC PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 CC PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
 CC PROSITE: PS00622; EGF_1; 3.
 CC PROSITE: PS01186; EGF_2; 2.
 CC Receptor, Tyrosine protein kinase, Transferase; Signal; ATP-binding;
 CC Repeat, EGF-like domain, Transmembrane; Immunoglobulin domain;
 CC Glycoprotein; Phosphorylation; Multigene family.
 CC SIGNAL: 1 18
 CC CHAIN: 19 1125 ANGIOPOIETIN 1 RECEPTOR
 CC DOMAIN: 19 746 EXTRACELLULAR (POTENTIAL)
 CC TRANSMEM: 747 771 POTENTIAL.
 CC DOMAIN: 772 1135 CYTOPLASMIC (POTENTIAL).
 CC DOMAIN: 44 102 IG-LIKE C2-TYPE DOMAIN 1.
 CC DOMAIN: 210 252 EGF-LIKE 1.
 CC DOMAIN: 254 296 EGF-LIKE 2.
 CC DOMAIN: 301 341 EGF-LIKE 3.
 CC DOMAIN: 370 424 IG-LIKE C2-TYPE DOMAIN 2.
 CC DOMAIN: 444 538 FIBRONECTIN TYPE-III 1.
 CC DOMAIN: 541 635 FIBRONECTIN TYPE-III 2.
 CC DOMAIN: 639 730 FIBRONECTIN TYPE-III 3.
 CC DOMAIN: 825 1097 PROTEIN KINASE.
 CC NP_BIND: 831 849 ATP (BY SIMILARITY).
 CC BINDING: 856 856 ATP (BY SIMILARITY).
 CC ACT_SITE: 965 965 BY SIMILARITY.
 CC MOD_RES: 993 993 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 CC CARBOHYD: 158 158 N-LINKED (GLCNAC...) (POTENTIAL).
 CC SEQUENCE: 1125 AA; 12502; MW: 615F132AR953B7F CR064.

Query Match 63.5%; Score 2458; DB 1; Length 1125.
 Best local Similarity 71.6%; Pred. No. 2.7e-163;
 Matches 471; Conservative 35; Mismatches 76; Indels 82; Gaps 7.

CY 1 MDSLASVLCVSVLLSGTVEGAKELILNSLFLVSDAEISLITIASWRPHEPIIGRD 60
 DB 1 MDSIAGIVLCCGVSTILSATVDGAMULLILNSLPLVSDAFTSLTICASWRPHEPIIGRD 60
 QY 61 FEALMNCQGHQDFVYQVNTFRWAKVWVKPKASKINCAIFCEGVPGEAIPITMMPQ 120
 EQ 61 FEALMNCQGHQDFVYQVNTFRWAKVWVKPKASKINCAIFCEGVPGEAIPITMMPQ 120
 QY 121 QASFLPATLITMTVIRKGNVNSFKKVLKREDAVLYKNGSFHSVPREHVDILLEVHLP 180
 DB 121 QASFLPATLITMTVIRKGNVNSFKKVLKREDAVLYKNGSFHSVPREHVDILLEVHLP 180
 QY 181 ADFQAGVYSARYIRINLITSAFTPLIVERFEAKWSEFENILELTA MNNAWCHETSEC 240
 DB 181 ADFQAGVYSARYIRINLITSAFTPLIVERFEAKWSEFENILELTA MNNAWCHETSEC 240
 QY 241 LCPDPMGRTTFAKTEIHTFRTCKERSGGGCKSVFCLPDYGGSCALQMKQCNF 300
 DB 241 LCPDPMGRTTFAKTEIHTFRTCKERSGGGCKSVFCLPDYGGSCALQMKQCNF 300
 QY 301 ACHPGYGPDCCKLRCSCHNNGEMCIHFQGLDASQWQGLQCEKESIPRMIPKIVDIPHE 360
 DB 301 ACHPGYGPDCCKLRCSCHNNGEMCIHFQGLDASQWQGLQCEKESIPRMIPKIVDIPHE 360
 QY 361 VNSKRFNPKICASGWPLTNEEMTLVKGEGIVLIPKDFNHTDHFSAITTHRLPLPUSG 420
 DB 361 VNSKRFNPKICASGWPLTNEEMTLVKGEGIVLIPKDFNHTDHFSAITTHRLPLPUSG 420
 QY 421 VVVCSTNTVAGVKEKPFNISVKVLPKPLNAPNVIDTGHFAVINISSPPYCEPKSCDKT 480
 DB 421 VVVCSTNTVAGVKEKPFNISVKVLPKPLNAPNVIDTGHFAVINISSPPYCEPKSCDKT 480
 QY 481 HTCPPCPAPELGGPSVFLPFPKPKDTLMISRTPTCVTVVWVSHDEPVEKVVYV 536
 DB 481 HTCPPCPAPELGGPSVFLPFPKPKDTLMISRTPTCVTVVWVSHDEPVEKVVYV 536
 QY 537 --DGEVEVINAKTPREFYVNSTYKVVSVLTVLHQLWLNKEYCKVSKNKAIPAEIKTIS 594
 DB 537 --DGEVEVINAKTPREFYVNSTYKVVSVLTVLHQLWLNKEYCKVSKNKAIPAEIKTIS 594
 QY 595 KAKGQPREPQVYTLPPSR--PEMTKNQVSIITCLVKGPYPS--DIATVWESNGQFENN 647
 DB 595 KAKGQPREPQVYTLPPSR--PEMTKNQVSIITCLVKGPYPS--DIATVWESNGQFENN 647
 QY 596 ASIS-----LPPKGLSLPKSQITLNLWLPFIPESSEDDFYVEVERKSVQMN 587
 DB 596 ASIS-----LPPKGLSLPKSQITLNLWLPFIPESSEDDFYVEVERKSVQMN 587
 RESULT 3
 ID TIE2_MOUSE STANDARD; PPT; 1122 AA.
 AC Q03858;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1995 (Rel. 31, last sequence update)
 DT 16-OCT-2001 (Rel. 40, last annotation update)
 DE Angiopoietin 1 receptor precursor (EC 2.7.1.12) (Tyrosine-protein
 DE kinase receptor TIE-2) (Tyrosine-protein kinase receptor TIE) (P140
 DE TEK) (Tunica interna endothelial cell kinase) (HYK).
 GN TEK OR TIE2 OR TIE-2 OR HYK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID:10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HAU/C; TISSUE=Lung;
 RX MEDLINE:9402374; PubMed:8415706;
 PA Sato T.N., Qin Y., Korak C.A., Andus K.L.;
 PT "Tie-1 and Tie-2 define another class of putative receptor tyrosine
 PT kinase genes expressed in early embryonic vascular system";
 RL Proc. Natl. Acad. Sci. U S A 90:9355-9358(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CD-1; TISSUE=Embryonic heart;
 RX MEDLINE:9324173; PubMed:3498682;
 PA Dumont D.J., Gradwohl G.T., Fouq G.-H., Auerbach E., Breitbart M.L.;

The endothelial-specific receptor tyrosine kinase, tek, is a member of a new subfamily of receptors.;
 Oncogene 8:1293-1301(1993).

SEQUENCE FROM N.A.
 TISSUE-Embryonic stem cells;
 MEDLINE-93129253; PubMed-1242811;
 Horita K., Yagi T., Kohmura N., Tomooka Y., Ikawa Y., Aizawa S.;
 "A novel tyrosine kinase, htk, expressed in murine embryonic stem cells.*;
 Biochem. Biophys. Res. Commun. 189:1747-1753(1992).

SEQUENCE FROM N.A.
 TISSUE-Lung;
 MEDLINE-94031116; PubMed-8217221;
 Kunking A.S., Stacker S.A., Wilks A.F.;
 "Tie2, a putative protein tyrosine kinase from a new class of cell surface receptor.*;
 Growth Factors 4:99-105(1993)

SEQUENCE FROM N.A.
 MEDLINE-9424447; PubMed-8187650;
 Schmeichel H., Risau W.;
 "Expression of tie-2, a member of a novel family of receptor tyrosine kinases, in the endothelial cell lineage.*;
 Development 119:957-968(1993).

SEQUENCE OF 822-1122 FROM N.A.
 STRAIN-CD-1; TISSUE-Embryonic heart;
 MEDLINE-92434855; PubMed-1630810;

Dumont D.J., Yamaguchi T.P., Conlon R.A., Rossant J., Breitman M.L.;
 "tek, a novel tyrosine kinase gene located on mouse chromosome 4, is expressed in endothelial cells and their presumptive precursors.*;
 Oncogene 7:1471-1480(1992).

FUNCTION: THIS PROTEIN IS A PROTEIN TYROSINE KINASE TRANSMEMBRANE RECEPTOR FOR ANGIOPOIETIN 1. IT MAY CONSTITUTE THE EARLIEST MAMMALIAN ENDOTHelial CELL LINEAGE MARKER. PROBABLY REGULATES ENDOTHELIAL CELL PROLIFERATION, DIFFERENTIATION AND GUIDES THE PROPER PATTERNING OF ENDOTHELIAL CELLS DURING BLOOD VESSEL FORMATION.

CATALYTIC ACTIVITY: ATP + a protein tyrosine -> ADP + protein tyrosine phosphate.

SUBCELLULAR LOCATION: Type I membrane protein

TISSUE SPECIFICITY: SPECIFICALLY EXPRESSED IN DEVELOPING VASCULAR ENDOTHELIAL CELLS.

DEVELOPMENTAL STAGE: EXPRESSION DETECTABLE IN DAY 8.5 EMBRYOS.

SIMILARITY: TO OTHER PROTEIN TYROSINE KINASES IN THE CATALYTIC DOMAIN.

SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-1/IFN-Gamma-TYPE II-DOMAINS.

SIMILARITY: CONTAINS 3 FIBRONECTIN TYPE III-LIKE DOMAINS.

SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See <http://www.isb.sib.ch/announcement/> or send an email to license@isb.sib.ch).

EMBL: X71426; CAA50557.1; F.

EMBL: X67553; CAA47857.1; F.

EMBL: D13748; BAA02883.1; F.

EMBL: S67051; AAB28663.1; F.

EMBL: S33142; S33142; F.

EMBL: M6198664; F.

EMBL: IPE000561; FGF-like

EMBL: IPE000719; FGF-like

EMBL: IPE003961; FN.III.

EMBL: IPE001245; IYI_EK_kinase.

EMBL: IPE00041; I03; F.

DR PRINTS: P000109; TYRKINASE.
 DR SMART: SM00181; EGF; 2.
 DR SMART: SM00060; FN3; 2.
 DR SMART: SM00219; TyKc; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYP; 1.
 DR PROSITE: PS00311; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00022; EGF_1; 3.
 DR PROSITE: PS01186; EGF_2; 3.
 KW RECEPTOR, tyrosine protein kinase, transferase; Signal: AIP-binding;
 KW Repeat, EGF-like domain; Transmembrane; Immunoglobulin domain;
 KW Glycoprotein; Phosphorylation; Multigene family;
 FT SIGNAL 1 18
 FT CHAIN 19 1122
 FT DOMAIN 19 744
 FT TRANSMEM 745 769
 FT DOMAIN 770 1122
 FT DOMAIN 44 102
 FT DOMAIN 210 252
 FT DOMAIN 254 299
 FT DOMAIN 301 341
 FT DOMAIN 370 424
 FT DOMAIN 444 536
 FT DOMAIN 539 633
 FT DOMAIN 637 731
 FT DOMAIN 822 1094
 FT NP_HIND 828 836
 FT BINDING 853 853
 FT ACT_SITE 962 962
 FT CARBOHYD 140 140
 FT CARBOHYD 158 158
 FT CARBOHYD 399 399
 FT CARBOHYD 438 438
 FT CARBOHYD 464 464
 FT CARBOHYD 558 558
 FT CARBOHYD 595 595
 FT CARBOHYD 648 648
 FT CARBOHYD 690 690
 FT MOD_RES 990 990
 FT CONFLICT 161 171
 FT CONFLICT 538 538
 FT CONFLICT 736 736
 FT CONFLICT 745 761
 FT CONFLICT 786 786
 FT CONFLICT 913 913
 FT CONFLICT 925 931
 FT CONFLICT 1117 1117
 SQ SEQUENCE 1122 AA; 125700 MW; F879630103FE96 CRC64;

Query Match 62.9%, Score 2432.5; DH 1; LOC431122;
 Best Local Similarity 68.1%; Pred No 1 60-161;
 Matches 481; Conservative 50; Mismatches 94; Indels 81; Gaps 14;
 QY 1 MESLASLVGVSLISIVGAMULLINSLELVSAETSLICIASWPHPEPTIGED 60
 DB 1 MESLASLVGVSLISIVGAMULLINSLELVSAETSLICIASWPHPEPTIGED 60
 QY 61 FEALMNHQDPLEVTQVTPREWAKVVKVVKPEKASKINFAVFCGVPVSEAIPIPTMPMPQ 120
 DB 61 FEALMNHQDPLEVTQVTPREWAKVVKVVKPEKASKINFAVFCGVPVSEAIPIPTMPMPQ 120
 QY 121 QASFLPATLTMTVDKGVNINISPKKVLKEDAVIYKNGSFTHSVPHHVVHLEVLPH 180
 DB 121 QASFLPATLTMTVDKGVNINISPKKVLKEDAVIYKNGSFTHSVPHHVVHLEVLPH 180
 QY 181 AGQTPARVYSAPYIGNIFTSAPTPLVPPCEAGKFWGPECNHLLTACNNWVWHEITGEC 240
 DB 181 AGQTPARVYSAPYIGNIFTSAPTPLVPPCEAGKFWGPECNHLLTACNNWVWHEITGEC 240
 QY 241 TCPPGFMGRTCEKACELHTFGRTCKERSQEGSKSVVPCLPDPYGGSTATGWKGLQNE 400
 DB 241 TCPPGFMGRTCEKACELHTFGRTCKERSQEGSKSVVPCLPDPYGGSTATGWKGLQNE 400


```

FT DISULFID 109 109 INTERCHAIN (WITH HEAVY CHAIN).
FT DISULFID 112 112 INTERCHAIN (WITH HEAVY CHAIN).
FT DISULFID 144 204
FT DISULFID 250 308
FT CARBOHYD 140 180
FT MOD_RES 330 330
FT VARIANT 97 97
FT VARIAT 259 259
FT VARIAT 241 241
FT STRAND 123 126
FT HELIX 130 134
FT TURN 136 137
FT STRAND 141 148
FT STRAND 158 162
FT TURN 163 164
FT STRAND 165 166
FT STRAND 175 178
FT STRAND 184 190
FT HELIX 193 197
FT TURN 198 199
FT STRAND 202 206
FT STRAND 215 219
FT STRAND 227 227
FT STRAND 240 234
FT HELIX 248 240
FT TURN 241 242
FT STRAND 245 256
FT STRAND 260 266
FT TURN 267 268
FT STRAND 259 270
FT STRAND 274 276
FT STRAND 280 281
FT TURN 283 284
FT STRAND 287 296
FT HELIX 297 301
FT TURN 302 303
FT STRAND 306 312
FT TURN 313 314
FT TURN 316 317
FT STRAND 320 324
SQ SEQUENCE 330 AA: 36106 MW: 37707P106C2P4330 C9C64.

Query Match 32.6%; Score 1262; DB 1; Length 330;
Best local Similarity 78.8%; Pred. No. 8,9c 81;
Matches 242; Conservative 12; Mismatches 11; Indels 42; Gaps 5.

QY 405 SVAITTHIRLP FDSGV WVCSTNTVAGMYEKFTNSV KVLPRKLNAPNVIDTG 457
Db 59 SSGLYSLSSVTVPSSSLSGTUTYICNVN-----HKPSNTRKDKKV----- 98
QY 458 HNFVAVINISSEHYFCEPKSCDKIHICPCAPPELLGGPSVFLPPPKKDILMSRPEVT 517
Db 99 -----EPKSCDKIHICPCAPPELLGGPSVFLPPPKKDILMSRPEVT 143
QY 518 CVVDVSHEDPEVKFNWYVDEVEVINAKTRPFFCYNSTYVWVSVLTLRQGLNLSKEYK 577
Db 144 CVVDVSHEDPEVKFNWYVDEVEVINAKTRPFFCYNSTYVWVSVLTLRQGLNLSKEYK 203
QY 578 QVSNKALPAPTEKTIKAKGAPPEPQVYTLPPSPREMTKNQVSLTLGVKGFPSDIAYE 637
Db 204 QVSNKALPAPTEKTIKAKGAPPEPQVYTLPPSPREMTKNQVSLTLGVKGFPSDIAYE 263
QY 638 WESNGQPNYNYKTPVLDSDGSEFLYSKLVIRKRWQCNQVISCVMHEALNHYIQKS 697
Db 264 WESNGQPNYNYKTPVLDSDGSEFLYSKLVIRKRWQCNQVISCVMHEALNHYIQKS 323
QY 698 LSLSPCK 704
Db 324 LSLSPCK 330

```

RESULT 5

```

GC3_HUMAN
ID GC3_HUMAN STANDARD: PRT; 290 AA.
AC P01860;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DI 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ig gamma-3 chain C region (heavy chain disease protein) (IHG).
GN IGHG3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE (DISEASE PROTEIN WIS).
RX MEDLINE=81021548; PubMed=6774747;
RA Frangione B., Rosenwasser E., Prelli F., Franklin E.C.;
RT "Primary structure of human gamma 3 immunoglobulin deletion mutant:
RT gamma 3 heavy-chain disease protein Wis.";
RL Biochemistry 19:4304-4308(1980).
RN [2]
RP REVISIONS TO 12-97 OF PROTEIN WIS.
RX MEDLINE=77118561; PubMed=402363;
RA Michaelisen T.E., Frangione B., Franklin E.C.;
RT "Primary structure of the 'hinge' region of human IgA3. Probable
RT quadruplication of a 15-amino acid residue basic unit.";
RL J. Biol. Chem. 252:883-889(1977).
RN [3]
RP REVISIONS TO 59-289 OF PROTEIN WIS (DISEASE PROTEIN ZUC).
RX MEDLINE=77021516; PubMed=823945;
RA Wolfenstein-Todel C., Frangione B., Prelli F., Franklin E.C.;
RT "The amino acid sequence of 'heavy chain disease' protein ZUC.
RT Structure of the Fc fragment of immunoglobulin G3.";
RL Biochem. Biophys. Res. Commun. 71:907-914(1976).
RN [4]
RP SEQUENCE FROM N.A. (DISEASE PROTEIN OMM).
RX MEDLINE=82247835; PubMed=6808505;
RA Alexander A., Steimetz M., Barritault D., Frangione B.,
RA Franklin E.C., Hood L., Buxbaum J.N.;
RT "Gamma Heavy chain disease in man: cDNA sequence supports partial
RT gene deletion model.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:3250-3264(1982).
CC -1- SUBUNIT: DIMER LINKED BY 12 DISULFIDE BONDS; IT HAS AN EXTRA
CC INTERCHAIN DISULFIDE BOND AT POSITION 7 IN ADDITION TO THE 11
CC NORMALLY PRESENT IN THE HINGE REGION.
CC -1- MISCELLANEOUS: THE HEAVY CHAIN DISULFIDE PROTEIN WIS IS SHOWN.
CC -1- MISCELLANEOUS: THE SEQUENCE OF RESIDUES 42-76 WAS TAKEN FROM THE
CC REF.2.
CC -1- MISCELLANEOUS: DISEASE PROTEIN WIS IS LACKING MOST OF THE V REGION
CC AND ALL OF THE CH1 REGION.
CC -1- MISCELLANEOUS: DISEASE PROTEIN ZUC LACK MOST OF THE V REGION, ALL
CC OF THE CH1 REGION, AND PART OF THE HINGE COMPARED WITH NORMAL.
CC GAMMA-3 HEAVY CHAINS.
CC -1- MISCELLANEOUS: DISULFIDE PROTEIN OMM MAY REPRESENT AN ALLELIC FORM
CC OR ANOTHER GAMMA CHAIN SUBCLASS.
CC -1- MISCELLANEOUS: THE HINGE REGION IN GAMMA-3 CHAINS IS ABOUT FOUR
CC TIMES AS LONG AS IN OTHER GAMMA CHAINS AND CONTAINS THREE
CC IDENTICAL 15-RESIDUE SEGMENTS PRECEDED BY A SIMILAR 17-RESIDUE
CC SEGMENT (12-28).
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J00231; AAA52805.1; ALT_SEQ.
DR PIR: A02149; G3H0W1.
DR HSPD: P01857; 1FC1.

```

```

DR MIM: 147120;
DR InterPro: IPR001006; Ig_MHC
DR InterPro: IPR001597; Ig_c1
DR InterPro: IPR003600; Ig_Like
DR Pfam: PF00047; Ig_2
DR SMART: SM00410; Ig_Like; 1
DR SMART: SM00407; IG_c1; 1
DR PROSITE: PS00290; IG_MHC; 1
KW Immunoglobulin domain; Immunoglobulin c region; glycoprotein; Repeat
FT DOMAIN 12 73
FT DOMAIN 74 183
FT DOMAIN 184 289
FT REPEAT 29 43
FT REPEAT 44 58
FT REPEAT 59 73
FT MOD_RES 1 1
FT MOD_RES 6 6
FT CARBOHYD 7 7
FT DISULFID 7 7
FT DISULFID 24 24
FT DISULFID 27 27
FT DISULFID 33 33
FT DISULFID 39 39
FT DISULFID 42 42
FT DISULFID 48 48
FT DISULFID 54 54
FT DISULFID 57 57
FT DISULFID 63 63
FT DISULFID 69 69
FT DISULFID 72 72
FT CARBOHYD 140 140
FT MOD_RES 290 290
FT VARIANT 126 127
FT VARIANT 134 134
FT VARIANT 139 139
FT VARIANT 182 182
FT VARIANT 227 227
FT VARIANT 227 227
FT VARIANT 279 279
FT VARIANT 279 279
FT SEQUENCE 290 AA, 326 AA, 32631 MW, 6646.457 kDa, pI 4.64.

Query Match 30.1%; Score 1163; DB 1; Length 290;
Best local Similarity 90.3%; Pred. No. 5,9e-74;
Matches 213; Conservative 11; Mismatches 12; Indels 0; Gaps 0,

QY 469 PYGPEPKSGKTHTPPCPPAPPELLAGPSVFLFPKPKCTLMISRTPEVTGVVVDVSHEDP 528
DB 1 ||||| || ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
55 PROPEPKSGDTPPCPPAPPELLAGPSVFLFPKPKCTLMISRTPEVTGVVVDVSHEDP 114
QY 529 EVKFNWVVKVVEVINAKTKPEFVYNSTYFVSVLTIVLHGLWLNGLKEYKCKVSNKALPAP 588
DB 1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
115 EVQKFWVDGVGVVINAKTKPEFQGNSTFRVSVLTIVLHQNWLDDGKEYCKVSNKALPAP 174
QY 589 IETKISKAKGQPREPQVYTLTPSPPEEMTKNQVSLTCLVKGFYPSDIAVEFESNQQPENNY 648
DB 1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
175 IETKISKTKQPREPQVYTLTPSPPEEMTKNQVSLTCLVKGFYPSDIAVEFESNQQPENNY 234
QY 643 KIIEPVLTSGSFYLSKILVKLSFKWLGQVNFNSVMHEALHNYITQKLSLSFGK 704
DB 1 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
235 NITPPMLDSGSPFLYSKLLIVDKSKWQGNIFSCVMHEALHNYITQKLSLSFGK 290

RESULT 6
GC2_HUMAN
ID GC2_HUMAN PRI: 326 AA.
AC P01859;

```

```

DT 21 JUL 1986 (Ref. 01, Created)
DT 21-JUL-1986 (Ref. 01, Last sequence update)
DT 16-OCT-2001 (Ref. 40, Last annotation update)
DE Ig gamma-2 chain C region.
GN IGHG2.
OS Homo sapiens (Human).
OC Eukaryota, Metazoa, Chordata, Vertebrata, Euteleostomi,
OC Mammalia, Eutheria, Primates, Catarrhini, Hominoidea, Homo.
OX NCRI_TaxID:9606;
RN [1]
RP SEQUENCE OF 2-326 FROM N.A.
RX MEDLINE=82197621; PubMed=6804948;
RA Ellison J.W., Hood L.E.;
RT "Linkage and sequence homology of two human immunoglobulin gamma
PL heavy chain constant region genes.";
PL Proc Natl Acad Sci U S A 79:1044-1048(1982).
PN [2]
RP SEQUENCE OF 88-115 FROM N.A.
RX MEDLINE=84001943; PubMed=6811139;
RA Takahashi N., Ueda S., Obara M., Nikaide T., Nakai S., Honjo T.;
RT "Structure of human immunoglobulin gamma genes: implications for
FT evolution of a gene family.";
FT Cell 29:671-679(1982).
RN [3]
RP SEQUENCE OF 99-177 AND 310-326 FROM N.A.
RX MEDLINE=84235992; PubMed=6306675;
RA Frawinkel N., Pabbitts T.H.;
RT "Comparison of the hinge coding segments in human immunoglobulin gamma
FT heavy chain genes and the linkage of the gamma 2 and gamma 4 subclass
FT genes.";
FT EMBO J. 1:403-407(1982).
RN [4]
RP SEQUENCE OF 1-325 (MYELOMA PROTEIN TIL).
RX MEDLINE=81007873; PubMed=6774012.
RA Wang A.C., Tung E., Fudenberg H.H.;
RT "The primary structure of a human IgG2 heavy chain: genetic,
FT evolutionary, and functional implications.";
FT J. Immunol. 125:1048-1054(1980).
RN [5]
RP SEQUENCE OF 1-85 AND 132-325 (MYELOMA PROTEIN ZIF).
RX MEDLINE=80001357; PubMed=113060;
RA Connell G.F., Parr D.M., Hofmann T.;
RT "The amino acid sequences of the three heavy chain constant region
FT domains of a human IgG2 myeloma protein.";
FT Can. J. Biochem. 57:758-767(1979).
RN [6]
RP SEQUENCE OF 238-275 (ZIF).
RX MEDLINE=80114419; PubMed=118920;
RA Hofmann T., Parr D.M.;
RT "A note of the amino acid sequence of residues 381-391 of human
FT immunoglobulin gamma chains.";
FT Mol. Immunol. 16:923-925(1979).
RN [7]
RP REVISIONS TO 25; 59; 60 AND 264-268 (ZIF).
RA Hofmann T., Parr D.M.;
RT Submitted (MAR 1980) to the PIR data bank.
RN [8]
RP SEQUENCE OF 1-121 (DOT).
RX MEDLINE=95255298; PubMed=7737190;
RA Stoppini M., Bellotti V., Negri A., Merlini G., Garver F., Ferri G.;
RT "Characterization of the two unique human anti-flavin monoclonal
FT immunoglobulins.";
FT Eur. J. Biochem. 224:886-893(1995).
RN [9]
RP DISULFIDE BONDS.
RX MEDLINE=72033500; PubMed=4940472;
RA Milstein C., Frangione B.;
RT "Disulphide bridges of the heavy chain of human immunoglobulin G2.";
FT Biochem. J. 121:217-225(1971).
RN [10]
RP DISULFIDE BONDS.

```

```

EX MEDLINE-64064124; PubMed-5782707;
RA Francione B., Milstein C., Pink J.R.L.;
RT "Structural studies of immunoglobulin G.";
RL Nature 221:145-148(1969).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J00230; AAB59393.1;
DR PIR: A02148; G2HU.
DR HSSP: P01857; 1FC1.
DR MIM: 147110;
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003597; Ig_C1.
DR InterPro: IPR003600; Ig_Like.
DR Pfam: PF00047; Ig_3.
DR SMART: SM00410; Ig_Like; 1.
DR SMART: SM00407; IgC1; 2.
DR PROSITE: PS00290; Ig_MHC; 2.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1
FT DOMAIN 1 98
FT DOMAIN 99 110
FT DOMAIN 111 219
FT DOMAIN 220 326
FT DOMAIN 327 347
FT DISULFID 14 14
FT DISULFID 27 83
FT DISULFID 102 102
FT DISULFID 103 103
FT DISULFID 106 106
FT DISULFID 109 109
FT DISULFID 140 200
FT DISULFID 246 304
FT SITE 156
FT MODRES 326 326
FT VARIANT 60 60
FT CONFLICT 109 109
FT SEQUENCE 326 AA; 35484 MW; 8416876567676767 Cn64;
Query Match 29.8%; Score 1154; DB 1; Length 326;
Best Local Similarity 82.4%; Pred. No. 2.8e-73;
Matches 215; Conservative 14; Mismatches 15; Indels 18; Gaps 2;
QY 461 AVINISSEPPFCEKSKSLKRNH-----LCPCHAPPELLGGPSVFLPDK 503
DB 67 SVVIVFSSNFGIQTICNDHDKPSNTKVDKIVKCKCCVECPCHPAPV-VAGPSVFLPDK 125
QY 504 PKDTLMISRTPEVTCVVVDVSHEDPEVKFNMYVDGSEVHNAKTPRDEQYNSTYRVVSVL 563
DB 126 PKDILMISRTPEVTCVVVDVSHEDPEVKFNMYVDGSEVHNAKTPRDEQYNSTYRVVSVL 185
QY 564 TVLHODWLNKKEYCKVSKNKAIPAPVTRKTIISKAKGQPRPEQVYTLPPSPREMTKNQVSLT 623
DB 186 TVVHODWLNKKEYCKVSKNKAIPAPVTRKTIISKAKGQPRPEQVYTLPPSPREMTKNQVSLT 245
QY 624 CLVKKCFYISDIADVWESNCGFENNYKTIPIVPLDSGDSFFLYSKITVDKSRWQOCNVPSCS 683
DB 246 CLVKKCFYISDIADVWESNCGFENNYKTIPIVPLDSGDSFFLYSKITVDKSRWQOCNVPSCS 305
QY 684 VMIEALHNHYTKSLSPK 704
DB 306 VMIEALHNHYTKSLSPK 326

```

RESULT 7
G4_HUMAN

```

ID GC4_HUMAN STANDARD; PRT; 327 AA.
AC P01861;
DT 21-JUL-1986 (rel. 01, Created)
DT 21-JUL-1986 (rel. 01, last sequence update)
DE 16-OCT-2001 (rel. 40, last annotation update)
DE Ig gamma-4 chain C region.
GN IGHS4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-83157104; PubMed-6299662;
RA Ellison J.W., Huxbaum J.N., Hood L.E.;
RT "Nucleotide sequence of a human immunoglobulin C gamma 4 gene.";
RL DNA 1:11-18(1981).
RN [2]
RP SEQUENCE OF 1-30 AND 81-326.
RX MEDLINE-70207560; PubMed-4192699;
RA Pink J.R.L., Huttery S.H., de Vries G.M., Milstein C.;
RT "Human immunoglobulin subclasses. Partial amino acid sequence of the
RT constant region of a gamma 4 chain.";
RL Biochem. J. 117:33-47(1970).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL: K01316; AAB59394.1; ALT_INIT.
PIR: A02150; G4HU.
HSSP: P01842; 7FAB.
MIM: 147130;
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR003597; Ig_C1.
DR InterPro: IPR003600; Ig_Like.
DR Pfam: PF00047; Ig_3.
DR SMART: SM00410; Ig_Like; 1.
DR SMART: SM00407; IgC1; 2.
DR PROSITE: PS00290; Ig_MHC; 2.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1
FT DOMAIN 1 98
FT DOMAIN 99 110
FT DOMAIN 111 220
FT DOMAIN 221 327
FT DISULFID 14 14
FT DISULFID 27 83
FT DISULFID 106 106
FT DISULFID 109 109
FT DISULFID 141 201
FT DISULFID 247 305
FT SEQUENCE 327 AA; 35940 MW; 3EDBD811EF208E7A Cn64;
Query Match 29.6%; Score 1144; DB 1; Length 327;
Best Local Similarity 89.9%; Pred. No. 1.4e-72;
Matches 213; Conservative 7; Mismatches 9; Indels 8; Gaps 1;
QY 468 EPPYCPKPSCKITHTCGCHAPPELLGGPSVFLPDKKTKTLMISRTPEVTCVVVDVSHED 527
DB 99 ESKYGGPP-----CFSPAPEELGGPSVFLPDKPKDTLMISRTPEVTCVVVDVSHED 150
QY 528 PEVKFNMYVDGSEVHNAKTPRDEQYNSTYRVVSVLTVLHQDWLNGKEYCKVSKNKAIPAPVTR 587
DB 151 PEVFNMYVDGSEVHNAKTPRDEQYNSTYRVVSVLTVLHQDWLNGKEYCKVSKNKAIPAPVTR 210
QY 588 PIETISKAKGQPRPEQVYTLPPSPREMTKNQVSLTCLVKGFYPSDIAVWESNCGFENNY 647

```


Db 211 SIERTISKAGQPRQVYTLPLPSQEMTKNAVSLTLVKKYFESDIAVWESNGSPERN 276
 QY 648 YKTIPTVLDSDGFFLYSKITVPSKWOQNVFSCSVMHAEALHNHYTKLSLSLSPCK 704
 DB 271 YKTIPTVLDSDGFFLYSKITVPSKWOQNVFSCSVMHAEALHNHYTKLSLSLSPCK 327

RESULT 8
 GC_HABIT STANDARD: PRT: 323 AA.
 AC P01870;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Ig gamma chain C region
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 FX MEDLINE=8403940; PubMed=6119520.
 RA Bernstein K F., Alexander C.R., Mage P.G.,
 RT "Nucleotide sequence of a rabbit IgG heavy chain from the recombinant
 RT F-1 haplotype".
 RL Immunogenetics 18:387-397(1983).
 RN [2]
 RP SEQUENCE OF 1-128.
 RX MEDLINE=76135469; PubMed=1243651;
 RA Pratt D.M., Mole L.E.;
 RT "Sequence studies on the constant region of the Fd sections of rabbit
 RT immunoglobulin G of different allotype".
 RL Biochem. J. 151:337-349(1975).
 RN [3]
 RP SEQUENCE OF 88-266 FROM N.A.
 RX MEDLINE=8329917; PubMed=3619512.
 RA Martens C.L., Moore K.W., Steinmetz M., Hood L., Knight K.L.;
 RT "Heavy chain genes of rabbit IgG: isolation of a cDNA encoding gamma
 RT heavy chain and identification of two genomic C gamma genes".
 RL Proc. Natl. Acad. Sci. U.S.A. 74:6019-6022(1982).
 RN [4]
 RP SEQUENCE OF 132-161.
 RX MEDLINE=70110015; PubMed=5461106;
 RA Fuchter P.G., Jackson S.A., Mole L.E., Porter P.R.;
 RT "Sequence studies of the Fd section of the heavy chain of rabbit
 RT immunoglobulin G".
 RL Biochem. J. 116:249-259(1970).
 RN [5]
 RP SEQUENCE OF 129-131 AND 155-122.
 RA Hill R.L., Lebovitz H.E., Fellows P.E. Jr., Delaney P.;
 RL (In) Killander J. (eds.);
 RL Gamma globulins. Nobel sym. 3, pp.109-127, Almqvist and Wiksell,
 RL Stockholm (1967).
 CC -1- MISCELLANEOUS: REF.1 SEQUENCE HAS THE D12 ALLOTYPIC MARKER,
 CC 104-THR, AND THE E14 MARKER, 185-THR. REF.3 HAS THE D11 AND E15
 CC MARKERS AND REF.5 THE E15 MARKER.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M16426, AAA1289.1,
 DR PIP, A02161, GHRP
 DR HSSP: P01857; 1FCL.
 DR InterPro: IPR003006; Ig_MRC.
 DR pfam: PF0001597; Ig_cl.
 DR SMART: SM00407; IgC1.2.
 DR PROSITE: PS00240; Ig_MHC; 1

Immunoglobulin family. Immunoglobulin C region.
 PW NON_TER 1
 FT VARIANT 104 104 T -> M (IN D11 MARKER).
 FT VARIANT 185 185 T -> A (IN E15 MARKER).
 FT CONFLICT 48 48 N -> E (IN REF. 2).
 FT CONFLICT 71 71 V -> VPV (IN REF. 2).
 FT CONFLICT 144 144 Q -> E (IN REF. 3 AND 4).
 FT CONFLICT 173 173 N -> D (IN REF. 5).
 FT CONFLICT 187 187 Q -> E (IN REF. 3 AND 5).
 FT CONFLICT 201 201 N -> D (IN REF. 5).
 FT CONFLICT 218 218 E -> Q (IN REF. 5).
 FT CONFLICT 232 232 E -> Q (IN REF. 5).
 FT CONFLICT 246 246 E -> G (IN REF. 5).
 FT CONFLICT 260 260 N -> D (IN REF. 5).
 FT CONFLICT 266 266 N -> D (IN REF. 5).
 FT CONFLICT 280 280 Y -> W (IN REF. 5).
 FT CONFLICT 284 284 N -> S (IN REF. 5).
 SQ SEQUENCE 323 AA: 69E8AA118D579A8B CRC64:
 Quality Match 24.1%, Score 332, pH 2, Length 323;
 Best Local Similarity 59.5%, Pct. No. 7.7e-58;
 Matches 173; Conservative 40; Mismatches 48; Indels 30; Gaps 3;
 QY 444 LPKPLNA-----PNVDTGHNFAVINISSEPYFGEPKSCDKTH----- 481
 DB 33 LPPEVTVWNSGTFNGVTFPSVRQSSGLYSLSVSVSSQPVTCNVAHPATNTRVD 92
 QY 482 -----TC--PPCAPELLGGPSVLEPPKPKDILMISRTPEVTGVVDVSHDEPKFN 533
 DB 93 KIVASTSTSKPLCPPEPELLGGPSVLEPPKPKDILMISRTPEVTGVVDVSHDEPKFN 152
 QY 534 WYVDGVEVHNAKTKRFEQYNSIYKVVSVLTVLHDMVNGKRYCKVSKNKAIPAIKTI 593
 DB 153 WYINQEVFAPPLPEPEQFNSTIPVSTFLPIHQLWELGFEFKCKVINKALPAIIEKTI 212
 QY 594 SKAKGQPEPQVYTLPLPSQEMTKNAVSLTLVKKYFESDIAVWESNGSPERNYKTTTP 653
 DB 213 SKAQGPPLPKVYTWGTPRELLSPSVSLTCMNGFYSDISVWEKRGKRAFDNKTTPA 272
 QY 654 VLDSDGFFLYSKITVPSKWOQNVFSCSVMHAEALHNHYTKLSLSLSPCK 704
 DB 273 VLDSDGFFLYSKITVPSKWOQNVFSCSVMHAEALHNHYTKLSLSLSPCK 323

RESULT 9
 GC2_CAVPO STANDARD: PRT: 329 AA.
 AC P01862;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Ig gamma-2 chain C region.
 OS Cavia porcellus (Guinea pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
 OX NCBI_TaxID=10141;
 RN [1]
 RP SEQUENCE OF 1-3.
 RA Trischmann T.M.;
 PL Submitted (APP-1975) to the PIP data bank.
 RN [2]
 RP SEQUENCE OF 4-68.
 RX MEDLINE=71058471; PubMed=5538606;
 FA Birshwin B.F., Hussain Q.Z., Cebra J.J.;
 ET "Structure of heavy chain from strain 13 guinea pig
 RT immunoglobulin-g(2). 3 Amino acid sequence of the region around the
 RL half-cysteine joining heavy and light chains.";
 RL Biochemistry 10:18-25(1971).
 RN [3]
 RP SEQUENCE OF 69-133 AND 312-329.
 RX MEDLINE=71058486; PubMed=5538616;

FT DOMAIN 783 1136 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 43 106 IG-LIKE C2-TYPE DOMAIN 1.
 FT DOMAIN 212 254 EGF-LIKE 1.
 FT DOMAIN 256 301 EGF-LIKE 2.
 FT DOMAIN 303 343 EGF-LIKE 3.
 FT DOMAIN 370 424 IG-LIKE C2-TYPE DOMAIN 2.
 FT DOMAIN 444 538 FIBRONECTIN TYPE-III 1.
 FT DOMAIN 541 637 FIBRONECTIN TYPE-III 2.
 FT DOMAIN 641 742 FIBRONECTIN TYPE-III 3.
 FT DOMAIN 837 1116 PROTEIN KINASE.
 FT NP_BIND 843 851 ATP (BY SIMILARITY).
 FT BINDING 868 868 ATP (BY SIMILARITY).
 FT ACT_SITE 977 977 HY SIMILARITY.
 FT MOD_RES 1005 1005 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT CARBOHYD 84 84 N-LINKED (GLCNAC) (POTENTIAL).
 FT CARBOHYD 159 159 N-LINKED (GLCNAC) (POTENTIAL).
 FT CARBOHYD 501 501 N-LINKED (GLCNAC) (POTENTIAL).
 FT CARBOHYD 594 594 N-LINKED (GLCNAC) (POTENTIAL).
 FT CARBOHYD 707 707 N-LINKED (GLCNAC) (POTENTIAL).
 SO SEQUENCE 1136 AA; 124953 MW; 8FF85804A041BB12 CRC64,
 Query Match 22.8%, Score 882.5, Dr 1, Length 1136,
 Best Local Similarity 33.5%, Pred. No. 9.8e-54,
 Matches 222; Conservative 89; Mismatches 250; Indels 101; Gaps 25;

Qy 16 LSGTVEGAMDLILNSPLVSDAETSLTCT-----ASGWRPEHPETIGKDFEALMN 66
 Db 15 LASHVGAADVLTLLADLRITPEORFFETLVSGEAGAGSGSDAWCP--PLLEKDDRTVPT 72

Qy 67 QH--QDPLEVTQVTRHWAKVVKWP--FKASKINGAYFCHGRVYRGEALRIETMKMQQAS 123
 Db 73 PRPWQPP-----HIARGSSKVTVRFSQPSLLGVSGEWSG-----GGGTRVLYVHNSPGAH 124

Qy 124 FLPATLTMDKGDNNVSEKVLLEEDAVIYKNGSFHSVPRHEVPD---ILEVHLPH 180
 Db 125 LLPDKVTHVTKCDTAVLSARVRKKEKTDVIWKSNGSYFYTLDRHEAQDQFLLQ--LPN 182

Qy 181 ACPDAGVYSAKYLIGNLTSAETLLIVPRCEAKWKSPENHLCTATMNNVSEHEDTGEC 240
 Db 183 VQPSGGIYSAIYLEASPLGSAFELLIVGCEAGWAGQCTKECPGCLHGAGVCHDQDEG 242

Qy 241 ICPHPCMCRTCEKACPLHTFRTCKEPCSGEPPKSVVPLPYPYSGVATKMKGLQENE 300
 Db 243 VCPGFTSTWTEATPSPSPQSDELPTASRSLPLPLPFPYSGSSSSSKWSALVE 302

Qy 301 ACHPCTGPGCKLEKSNHMDLPFGNLTSPWATGLVTEREHIIPMTPKIIVLDPHIE 360
 Db 303 ACAPGRFCAVCHIQCCQNGCTGTATFSGCVCPSSWHCMHCKKSI---PLIQILLDMVSELE 359

Qy 361 VNSCKENPI--CKASGWPPLTNEENTLVKPGSTV--LHPKENHILHIESVALETHIRILPPD 418
 Db 360 FNLDTPRINCAGNPFVYSGSMELRKPSTVLLSTKAIVEPDP--TTAFEFVPRFALGD 418

Qy 419 SGVWVCVNTVAGVMPKPFNIISKVLKPLNAPNVDVTHGNFAVINLSRPYFCEPKSCD 478
 Db 419 SGLWECRVSTSGGDSRPFPINVKVPVPVPIAPLLAKQSPQLVVS-----464

Qy 479 KTHICPPCPAPELIG--CP--SVPLFPKPDKDTLMSPTPTVTCVVVVSIE-----DPE 529
 Db 465 -----PLVPSGDPPIASVRIH--YRPQDSTMAWST-----IVVDPSENVTLIMLRPK 510

Qy 530 VKFNWYVGVVEVINAKTKPRFELVNS-----YRWVSVLVLHQLWLNKEY-----KCKVS 581
 Db 511 T-----GYSVRVQLSRPEEGEGNAGWPPFLMTITPEPLLPKWLGEWIVEFSPDLFVS 564

Qy 582 NKALPAP-----IEKTSKACQPPRPQCVYTIHPSRFPMIKNCSVFQIVKSG--FYYS 632
 Db 564 WSLPVPVPGPLVGQGTFLRLWCAAPQERRENV--SSPVAFTAL-----LTSLTPGTGYQL 616

Qy 633 DI 634
 Db 617 DV 618

RESULT 11

TIE1_HUMAN

ID TIE1_HUMAN STANDARD: PRT: 1138 AA.
 AC P35590;
 DT 01-JUN-1994 (rel. 29, Created)
 DT 01-JUN-1994 (rel. 29, Last sequence update)
 DT 15-JUL-1996 (rel. 38, Last annotation update)
 DE Tyrosine-protein kinase receptor TIE-1 precursor (EC 2.7.1.112).
 GN TIE1 OR TIE.
 OS Homo sapiens (Human).
 OC Eukaryota, Metazoa, Chordata; Craniata, Vertebrata, Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID:9606;
 RN [1]
 PP SEQUENCE FROM N.A.
 RX MEDLINE:92195316; PubMed-1312667;
 RA Partanen J., Armstrong E., Mackelae T.P., Korhonen J., Sandberg M.,
 RA Renkonen R., Knuutila S., Huebner K., Alitalo K.;
 RT "A novel endothelial cell surface receptor tyrosine kinase with
 extracellular epidermal growth factor homology domains.";
 FL Mol Cell Biol 12 1638-1707(1992).
 RN [2]
 RP REVISIONS.
 RA Partanen J.M.;
 RL Submitted (JUL-1993) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: PROBABLE PROTEIN TYROSINE-KINASE TRANSMEMBRANE RECEPTOR.
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -!- SUBCELLULAR LOCATION: Type 1 membrane protein.
 CC -!- TISSUE SPECIFICITY: SPECIFICALLY EXPRESSED IN DEVELOPING VASCULAR
 CC ENDOTHELIAL CELLS.
 CC -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
 CC DOMAIN.
 CC -!- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC -!- SIMILARITY: CONTAINS 3 FIBRONECTIN TYPE III-LIKE DOMAINS.
 CC -!- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.
 CC
 CC This SWISS-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC FMBL: X60957; CAA43290.1;
 CC HSSP: P11362; 1PGK.
 CC MIM: 609223;
 DR InterPro: IPR000561; EGF-like
 DR InterPro: IPR000719; Euk_kinase.
 DR InterPro: IPR000461; FN_III.
 DR InterPro: IPR003006; Ig_MHC.
 DR InterPro: IPR01245; Tyr_kinase.
 DR Pfam: PF00008; EGF_2.
 DR Pfam: PF00041; In3; 3.
 DR Pfam: PF00047; Ig; 2.
 DR Pfam: PF00664; fkinase; 1.
 DR PRINTS: PR00109; TYRKINASE.
 DR SMART: SM00181; PGP; 2.
 DR SMART: SM00060; FN3; 2.
 DR SMART: SM00219; TyrKc; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYP; 1.
 DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00022; EGF_1; 3.
 DR PROSITE: PS01186; EGF_2; 3.
 KW Receptor; Tyrosine-protein kinase; Transferase; Signal; ATP-binding;
 KW Repeat; EGF-like domain; Transmembrane; Immunoglobulin domain;
 KW Glycoprotein; Phosphorylation; Multigene family.
 FT SIGNAL 1 24 POTENTIAL.
 FT CHAIN 25 1138 TYROSINE-PROTEIN KINASE RECEPTOR TIE-1.


```

Db 269 LEODYKNTPTILDSGTYFYLSKLTIVTDSWLOQETFTCSVWHEALHNHHHTKNLSRSPG 328
QY 704 K 704
Db 329 K 329

RESULT 14
GCB_RAT
ID GCB_RAT STANDARD: PRT: 333 AA.
AC P20761:
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15 JUL 1999 (Rel. 38, Last annotation update)
DE Iq gamma-2B chain C region.
OS Rattus norvegicus (Rat).
OC Eukaryota, Metazoa, Chordata, Vertebrata, Euteleostomi;
OC Mammalia, Eutheria, Rodentia, Sciurognathi, Muridae, Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89232738; PubMed=3149946;
RA Bruegemann M.;
RT "Evolution of the rat immunoglobulin gamma heavy-chain gene family.";
RL Gene 74:473-482(1988).
DR PIR: PS0018; PS0018.
DR HSSP: P01842; 7FAB.
DR InterPro: IPR003006; Iq_MHC.
DR InterPro: IPR003597; Iq_cl.
DR Pfam: PF00047; Iq; 3.
DR SMART: SM00410; IG_Like; 1.
DR SMART: SM00407; IGcl; 2.
DR PROSITE: PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region.
FT NON_TER 1 1
FT DISULFID 15 15 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 27 80
FT DISULFID 106 106 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 112 112 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 115 115 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 147 207
FT DISULFID 253 311
SQ SEQUENCE 333 AA: 3649/ MW: 55F8664046L460A6 CR264;

Query Match
Best local Similarity 43.4%; Pred. No. 3,2e-52;
Matches 163, Conservative 47, Mismatches 73, Indels 47, Gaps 3;

QY 415 LPPDSGVWCVSNTVAGMVEKPNISVKVLPKPLNA-----PNVIDTGHNF 460
Db 11 LAOCGGTSTSTVTGLVAGYF-----PEPVTVTWSGALSSDVHTFPAVLQSLYR 63
QY 461 AVINISSEPFGEKPKSDKTH-----TCPPCPAPELIGG 494
Db 64 LTSVSTSTSPSQVTNVAHPASSTKVDKKVERPNNGIHKPTPTCHKCPVPELIGG 123
QY 495 PSVFLFPKPKDITIMISRTPEVTCVVVDVSDHEDPEVKFNMYVDQVFNHAKTKPREFOYN 554
Db 124 PSVFIFFPKPKDILLISQNAKVICVVVDVSEEPQVQFSWVNVEVHTAQTUPPEQYN 184
QY 555 STYRVSVLVLHODWLNKGEYKCKVSNKALPAPIEKTISKAKGQPREPOVYTLPPSRRE 614
Db 184 STFFVVSALPIQHDWMSGKEFKCKVNNKALPSPIETISKPKCLVKIKVQVYVWGPTPEQ 243
QY 615 MTKNQVSLTCLVKGFYPSDIAVEESGQPENNYKTPPPVLDSDGSFFLYSKLTVDKSRW 674
Db 244 LTEGVTSLTCLTSGLFENDLGVEWTSNGLHIEKKNKNTPEPVWDSGGSFFMTSKLNVSRW 303
QY 675 QQGNVSCSVNHEALHNHYTQKSLSISPGK 704

```

```

Db 304 DSRAPFVCSVWHEGLNHHVKSISNPPCK 333
QY 409 FTTHPLTPDPSGVW-----VCSVNTVAGMVEKPNISVKVLPKPLNAPWITGHNFAVI 463
Db 61 YSLSSIVTVPSSTWPSQVTICNVNVAHPASSTKVDKKVERPNNGIHKPTPTCHKCPVPELIGG 108

RESULT 14
GCB_MOUSE
ID GCB_MOUSE STANDARD: PRT: 398 AA.
AC P03987:
DT 23-OCT-1986 (Rel. 02, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15 JUL 1999 (Rel. 38, Last annotation update)
DE Iq gamma-3 chain C region, membrane-bound form.
OS Mus musculus (Mouse).
OC Eukaryota, Metazoa, Chordata, Vertebrata, Euteleostomi;
OC Mammalia, Eutheria, Rodentia, Sciurognathi, Muridae, Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85027161; PubMed=6392053;
RA Wells J.A., Ward C.J., Pimm D., Der-Halan G.P., Martinez H.M.,
RA Tucker P.W., Blattner F.R.;
RT "Structural analysis of the murine IgG3 constant region gene.";
PL EMBO J. 3:2041-2046(1984).
RN [2]
RP SEQUENCE OF 328-398 FROM N.A.
RX MEDLINE=84041483; PubMed=6314258;
RA Komaromy M., Clayton L., Rogers T., Robertson S., Kettman J.,
RA Wall R.;
RT "The structure of the mouse immunoglobulin in gamma 3 membrane gene segment.";
RL Nucleic Acids Res. 11:6775-6785(1983).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC EMBL: J00451; AAB59655.1; .
CC PIR: V01526; CAA24767.1; AIT_SFQ.
CC PIR: A02155; G3MSM.
CC HSSP: P01857; 1FCL.
CC InterPro: IPR003006; Iq_MHC.
CC InterPro: IPR003597; Iq_cl.
CC InterPro: IPR003600; Iq_Like.
CC Pfam: PF00047; Iq; 3.
CC SMART: SM00410; IG_Like; 1.
CC SMART: SM00407; IGcl; 2.
CC PROSITE: PS00290; IG_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein; Transmembrane; Alternative splicing.
FT NON_TER 1 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 113 HINGE.
FT DOMAIN 114 233 CH2.
FT DOMAIN 234 327 CH3.
FT TRANSMEM 346 362 POTENTIAL.
FT DOMAIN 363 398 CYTOPLASMIC (POTENTIAL).
FT CONFLICT 333 343 E -> G (IN REF. 2).
FT CONFLICT 342 342 E -> Q (IN REF. 2).
FT CONFLICT 388 388 P -> F (IN REF. 2).
SQ SEQUENCE 398 AA: 43929 MW: C7F7264B50A41B95 CRK64;

Query Match
Best local Similarity 21.9%; Score 84e-5; DB 1; Length 398;
Matches 159, Conservative 41, Mismatches 62, Indels 37, Gaps 3;

QY 409 FTTHPLTPDPSGVW-----VCSVNTVAGMVEKPNISVKVLPKPLNAPWITGHNFAVI 463
Db 61 YSLSSIVTVPSSTWPSQVTICNVNVAHPASSTKVDKKVERPNNGIHKPTPTCHKCPVPELIGG 108

```

QY 464 NISSEYFGEFKSCDK PHTCQDPGPAPELLGQSVLEFPKPKDTLMISRTPEVTCVVVDV 523
DB 109 -----SSCPGNTILGSPSVFIPFPKPKDALMSLTPKVTCTVVDV 148
QY 524 SHEDPEKFNKYVGVVYVHNNAIKRPRKYNSTYKVSVALIVHQWLNCKEYKCKVSNK 583
DB 149 SFDDPDVHVSVFVNKEVHTAWIQHRAQYNSTFRVVSATPIQHQIWMKPKFKCKVNNK 208
QY 584 ALPAPIEKTISKAKQPREPAVYITLPPSRREMTKNQVSLTCLKGFEYPSDIAVWESNGQ 643
DB 209 ALPAPIERTISKPKRAGQAVYITLPPSRREMTKNQVSLTCLKGFEYPSDIAVWESNGQ 208
QY 644 PENNYKITTTPVLDSGSEFLYSKLTIVUKSRWGGCNVSCSMHPRALNNHYTKRSLSISP 702
DB 265 LEQYKNTPTPLDSTSTFLYSKLTIVUKSRWGGCNVSCSMHPRALNNHYTKRSLSISP 427
RESULT 15
GCL_MOUSE STANDARD: PRT: 324 AA.
AC P01868;
DT 21-JUL-1986 (Rel. 01, Created)
DI 21-JUL-1986 (Rel. 01, Last sequence update)
DI 30-MAY-2000 (Rel. 39, Last annotation update)
DE Ig gamma-1 chain C region.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID:10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-80045036; PubMed-115593;
RA Honjo T., Ohta M., Yamawaki-Kataoka Y., Kataoka T., Kawakami T.,
RA Takahashi N., Mano Y.
RT Cloning and complete nucleotide sequence of mouse immunoglobulin
RT gamma 1 chain gene.
RL Cell 18:559-568 (1979).
RN [2]
RP SEQUENCE OF 76-324 FROM N.A. (MYELOMA PROTEIN MOPC 31C).
RX MEDLINE-80202559; PubMed-676752;
RA Ohta M., Yamawaki-Kataoka Y., Takahashi N., Kataoka T., Shimizu A.,
RA Mano Y., Seidman J.G., Peterlin H.M., Leder P., Honjo T.
RT Immunoglobulin gamma 1 heavy chain gene: structural gene sequences
RT cloned in a bacterial plasmid.
RL Gene 9:87-97 (1980).
RN [3]
RP SEQUENCE OF 7C-322 FROM N.A. (MYELOMA PROTEIN MOPC 21).
RX MEDLINE-80012837; PubMed-113776;
RA Rogers J., Clarke P., Salser W.
RT Sequence analysis of cloned cDNA encoding part of an immunoglobulin
RT heavy chain.
RL Nucleic Acids Res. 6:3305-3321 (1979).
RN [4]
RP SEQUENCE (MYELOMA PROTEIN MOPC 21).
RX MEDLINE-78242288; PubMed-48524;
RA Adelman K.
RT Evolution of immunoglobulin subclasses. Primary structure of a
RT murine myeloma gamma1 chain.
RL J. Biol. Chem. 253:6068-6075 (1978).
RN [5]
RP DISULFIDE BONDS (MOPC 21).
RX MEDLINE-74088889; PubMed-5673237;
RA Svasti J., Milstein C.
RT The disulphide bridges of a mouse immunoglobulin G1 protein.
RL Biochem. J. 126:837-850 (1972).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to license@sib-sib.ch.

DB EMBL: V00793; CAA24172.1; -
DB EMBL: V00793; CAA24173.1; -
DB EMBL: V00793; CAA24174.1; -
DB EMBL: V00793; CAA24175.1; -
DB EMBL: V00795; CAA24176.1; -
DB PIR: A02159; GINS.
DB HSSP: P01842; 7FAB.
DB GLYCOSULEDB: P01868; -
DB MGD: MGI:96446; Igh-4.
DB InterPro: IPR003006; Iq_MHC.
DB InterPro: IPR003597; Iq_C1.
DB Pfam: PF00047; Iq_3.
DB SMART: SM00407; IqC1_2.
DB ProSITE: PS00290; Iq_MHC; 1.
KW Immunoglobulin domain; Immunoglobulin C region; Glycoprotein;
KW Alternative splicing.
FT NON_TER 1
FT DOMAIN 1 97 CH1.
FT DOMAIN 98 110 HINGE.
FT DOMAIN 111 217 CH2.
FT DOMAIN 218 324 CH3.
FT DISULFID 27 82 INTERCHAIN (WITH A LIGHT CHAIN).
FT DISULFID 102 102 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 104 104 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 107 107 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 109 109 INTERCHAIN (WITH A HEAVY CHAIN).
FT DISULFID 138 198 N-LINKED (GLCNAC...).
FT CARBOHYD 174 174 /FTID-CAR_000055.
FT DISULFID 244 302
FT MOD_RES 324 324 REMOVED POST-TRANSLATIONALLY.
FT CONFLICT 276 276 N -> D (IN REF. 3).
FT CONFLICT 278 278 N -> D (IN REF. 3).
SQ SEQUENCE 324 AA; 35704 MW; A38612F3D1F2C93 CRC64;

Query Match 21.8% Score 842; DB 1; Length 324;
Best Local Similarity 49.1%; Pred. No. 1.4e-51;
Matches 159, Conservative 56, Mismatches 65, Indels 44, Gaps 6;

QY 415 LPPDSGVWVCVNIVACMKVPEFNI SVKVIKPL-----NAPNVIDYCHNF----- 460
DB 11 LAPGSAAVINSMVILGCLVKGYP-----PEPVTVTNSSGSLSSGVHTFPVAVLSQSLYT 63
QY 461 --AVINISSEYFGEKSCDKTHT-----CPP--CPAPPELLGSPSWFLF 500
DB 64 LSSSVTVPSRPSPKSTVTCNVAHPASSITKIVKIPVDCGCKPCICIVPEV---SSVFI 120
QY 501 PPKPKDTLMTSPTPEVTCVVVTVVSHEDPEKFNKYVGVVYVHNNAIKRPRKYNSTYKVS 540
DB 121 PPKPKDVLITITPKVTCTVVDVDSKDDPEVQFSWFDVDDVEVHTAQTQPREEQNFSTFRV 180
QY 561 SVLVIVIHQDWLNKPKYKCKVSNKALPAPIEKTISKAKQPREPAVYITLPPSRREMTKNQV 620
DB 181 SELPMIHQDWLNKPKYKCKVSNKALPAPIEKTISKAKQPREPAVYITLPPSRREMTKNQV 240
QY 621 SLTCLVKGFEYPSDIAVWESNGQFPENNYKITTTPVLDSGSEFLYSKLTIVUKSRWGGCNV 680
DB 241 SLTCLMTIDFPEDITVEQWQNCQPAENYKNTQPIMNINGSYFVYSKINUKSNWACNT 300
QY 681 SCVNMIEALNNHYTKRSLSISP 704
DB 301 TCSVLIEGLNHHHTKRSLSISP 324

Search completed: September 23, 2002, 22:13:46
Job time: 197 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 23, 2002, 22:09:44 Search time: 72.16 seconds
(without alignments)
1687.755 Million cell updates/sec

Title: US-09-733-764-2
Perfect score: 3069
Sequence: 1 MDSLASLVGVSVLLSGTV MIEALINHYTKSLSPGK 704

Scoring table: BLOSUM62 Gapop 10.0, Gapext 0.5

Searched: 56222 seqs, 17299429 residues

Total number of hits satisfying chosen parameters: 56222

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: SPTRMBL19:

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacterioph.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance 'n' have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2293.5	59.3	1082	11 Q9W24	Q9W24 rat
2	1303	33.7	701	4 Q9W24	Q9W24 rat
3	916.5	23.7	337	6 Q9W24	Q9W24 rat
4	901.5	23.3	1116	13 Q9W24	Q9W24 rat
5	849	21.9	463	11 Q9W24	Q9W24 rat
6	845	21.8	437	11 Q9W24	Q9W24 rat
7	825	21.3	473	11 Q9W24	Q9W24 rat
8	817	21.1	473	11 Q9W24	Q9W24 rat
9	816	21.1	468	11 Q9W24	Q9W24 rat
10	788	20.4	473	11 Q9W24	Q9W24 rat
11	360	9.3	375	4 Q9W24	Q9W24 rat
12	359	9.3	375	4 Q9W24	Q9W24 rat
13	359	9.3	375	4 Q9W24	Q9W24 rat
14	359	9.3	375	4 Q9W24	Q9W24 rat
15	355	9.2	613	4 Q9W24	Q9W24 rat
16	349	9.0	614	4 Q9W24	Q9W24 rat

ALIGNMENTS

RESULT 1

ID	Q9W24	PRELIMINARY	PRT	1083 AA
AC	Q9W24			
BT	01-MAY-2000 (TRIMBLrel. 13, Created)			
DT	01-MAY-2000 (TRIMBLrel. 13, Last sequence update)			
DT	01-DEC-2001 (TRIMBLrel. 19, Last annotation update)			
DE	TIE-2 RECEPTOR-LIKE TYROSINE KINASE			
OS	Rattus sp.			
OC	Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi;			
OC	Mammalia, Eutheria, Rodentia, Sciurognathi, Muridae, Murinae, Rattus.			
OX	NCHI_TaxID=10118;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
PX	MEDLINE:93275658; PubMed-7684830;			
PA	Maisonnier P.C., Goldfarb M., Yancopoulos G.D., Gao G.			
KT	"Distinct rat genes with related profiles of expression define a TIE			
RT	receptor tyrosine kinase family."			
RL	Oncogene 8:1631-1637(1993).			
DR	HSSP: P11362; IFGK.			
DR	InterPro: IPR000561; EGF-like.			
DR	InterPro: IPR000719; Euk_pkinase.			
DR	InterPro: IPR003961; FN_III.			
DR	InterPro: IPR001245; Tyr_pkinase.			
DR	InterPro: IPR00041; In3; 3.			
DR	PRINTS: PF00069; pkinase; 1.			
DR	PRINTS: PF00011; EGF_LAMININ.			
DR	SMART: SM00109; TYRKINASE.			
DR	SMART: SM00180; EGF_Lam; 1.			
DR	SMART: SM00001; EGF_Like; 1.			
DR	SMART: SM00239; FN3; 3.			
DR	PROSITE: PS00022; EGF_1; ENKRN_3.			
DR	PROSITE: PS01186; EGF_2; 3.			
DR	PROSITE: PS0107; PROTEIN_KINASE_ATP; 1.			
DR	PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.			
DR	PROSITE: PS0119; PROTEIN_KINASE_TYP; 1.			
KW	ATP-binding; EGF like domain; Glycoprotein; Transferase;			

KW Tyrosine-protein kinase.

SQ SEQUENCE 1081 AA: 122024 MW: 649750 GF94A30 7664.

Query Match 59.3%; Score 2293.5; DH 11; Length 1083;
 Best Local Similarity 68.0%; Pred. No. 1.7e-202;
 Matches 453; Conservative 40; Mismatches 96; Indels 77; Gaps 13;

QY 40 TSITCISASQWHPHPIICRQFEALMNGHQDLEVTQVIFREWAKKVVWKKRASKINGA 99
 DB 1 TSITCISASQWHPHPIICRQFEALMNGHQDLEVTQVIFREWAKKVVWKKRASKINGA 60
 QY 100 YFEIPIVRGEGATIRITMKMRQASLPAILTMTVDKIDNNNISPKVLIKEEDAVIYKNG 159
 DB 61 YFEIPIVRGEGATIRITMKMRQASLPAILTMTVDKIDNNNISPKVLIKEEDAVIYKNG 120
 QY 160 SFTHSPREHVPDILEVHLPHAQDQACVYSARYIGCNIFTSATRIIVRCAQKQPE 219
 DB 121 SFTHSPREHVPDILEVHLPHAQDQACVYSARYIGCNIFTSATRIIVRCAQKQWPD 180
 QY 220 CNHQAIFACMNNCGVCHEDIEGICITGPHMERICEKACIEHIFGRICKEKRCGSGQCKSYVF 279
 DB 181 CNHQAIFACMNNCGVCHEDIEGICITGPHMERICEKACIEHIFGRICKEKRCGSGQCKSYVF 240
 QY 280 CLPDYVGSATSWKGLTQNFANRHFYSPKYLEGSCNNNEMIDREDAKGLASFGWGLQ 349
 DB 241 CLPDYVGSATSWKGLTQNFANRHFYSPKYLEGSCNNNEMIDREDAKGLASFGWGLQ 300
 QY 340 CRREGIPKMTPIKIVLDPHIEVNSCKFNICKASGWIPTNEEMTLVKPDGTVLIRPKDFN 399
 DB 301 CRREGIPKMTPIKIVLDPHIEVNSCKFNICKASGWIPTNEEMTLVKPDGTVLIRPKDFN 360
 QY 400 HHDHF-SVAIFTHRIIPDSDVWVCSVNTVACMVKEKPNISVKVLPKIDINAPNVIDTGHN 459
 DB 361 HHDHF-SVAIFTHRIIPDSDVWVCSVNTVACMVKEKPNISVKVLPKIDINAPNVIDTGHN 420
 QY 460 FAVINISSEPYRGE-PKSCDKTHTCP----- 484
 DB 421 FAVINISSEPYRGE-PKSCDKTHTCP----- 484
 QY 485 PCPAPALHAGP-----SVFETPKPKDTELMISRTD-EVTCVVVDVSHRDPVKFNWYV 536
 DB 481 PCGEGEGHPGVRERTTASIGLPPPPGSLILPKSQTALNLTQPIETSSD-----EFY- 534
 QY 537 DGVVEVHNAKIKREHQYNSITRVSVSLT-VLHQWLKCKEYKCKVSNKALPAPIEXTISK 595
 DB 535 --VEVERNSQOTRSQOON--IKVPGNLTISVILNNLIPRQYSVRA-----RVNTK 580
 QY 596 AKGQ-PRPQVY-----LPPSBEEM-TRNOVSLTCLVK-----GFYPSDIAVEWESNG 642
 DB 581 ACGWSEERAWTLDSDSLPPGPNIKITINIDYALVSVTIVDGYSSISIIIRYKVGQ 640
 QY 643 QPENNY 648
 DB 641 KNEQDQ 646

RESULT 2

ID Q96PQ8 PRELIMINARY; PRT: 701 AA.
 AC Q96PQ8;
 DT 01-DEC-2001 (Tremblrel, 19, created)
 DT 01-DEC-2001 (Tremblrel, 19, last sequence update)
 DT 01-DEC-2001 (Tremblrel, 19, last annotation update)
 DE FACTOR VII ACTIVE SITE MUTANT IMMUNOCONJUGATE.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN 11
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21477448; PubMed=11593034;
 RA Hu Z., Garen A.;

RT Targeting tissue factor on tumor vascular endothelial cells and tumor
 FI cells for immunotherapy in mouse models of prestatial cancer*;
 RU Proc. Natl. Acad. Sci. U.S.A. 98:12180-12185(2001).
 DB EMBL: AF272774; AAK58686.1;
 SQ SEQUENCE 701 AA: 77826 MW: 94AC6CHM42C992F CRC64;

Query Match 33.7%; Score 1303; DB 4; Length 701;

Best Local Similarity 46.6%; Pred. No. 2.6e-111;

Matches 307; Conservative 36; Mismatches 112; Indels 204; Gaps 26;

QY 205 PLIVPEF-TAQK--W-----RPFNNHILATA'NNN'WVHEDIGS---CIQPPPIPMQPTTEK 253
 DB 88 REIFDAERIKIPWTSYSDRQK---ASSP'QNGSGKQIQLSYICFCLPAFEGPNTETIK 145
 QY 254 ---ACELHTFCRICKPCRCGOGCKSVFCLPDIYGCSCATGKGIQCNKACHPGHYCP 309
 DB 146 EDLLICVNNNGS--CEQY'SCHIGIKP-----SCKHEGYSLLAGVSCITPIVEYP 194
 QY 310 DQ-----KLMS'NNNGEMTQFQWTL'SPR---WAGLAF'ERFGIPMTPIKIVDLPDI 359
 DB 195 -CGKIPILKKRNASKPQGRIVG---GKVCIPKGFCHQVLI----- 230
 QY 360 EVNSCK-----FNPI-----C--KASGW---LPTNEEMTLVKPDG-----TVLH 394
 DB 241 LVNQAQAGRIILINTIIVVSSAHP'FDKTKNPNILIAVAGFILL'SEHDTEQSPFAVAJ 290
 QY 395 PKDF--NHDPFSAVETIHR-LLPDISGVVWCVSNIVACMVKEKPN----- 448
 DB 291 PSTYVPGSTNH-DIALLEHQPVVLTIRVVPLG-----LPEFTSEKILAVRSLVS 442
 QY 439 -----ISVKVLKPL-----NAPNVID----- 455
 DB 343 GWCQILDKGATALEIMVINVPRIMEQCIQGSRAKVGSPNITFYMPCAGYSNGSKNSCA 402
 QY 456 -----TGHNFAVINS-----SEYFG----- 472
 DB 403 DSGGPIATHYRGTVLTGIVSWGQCATVGHGVYTVRSQYVIEWLQKMRSEPRPVLLR 462
 QY 473 -----EPKCDKTHICPPAPPELLGGPSVFLPFPKPKDTELMISRTDQVTVVSH 525
 DB 463 APFPGSAFPKSKDTHCTPP'PAPELLGSPSVFLPFPKPKDTELMISRTDQVTVVSH 522
 QY 526 EDPEVKFNWYVGVVEVINAKTKPREEQYNSTYRVSVSLVLIHQWLKCKEYKCKVSNKAL 585
 DB 523 EDPEVKFNWYVGVVEVINAKTKPREEQYNSTYRVSVSLVLIHQWLKCKEYKCKVSNKAL 582
 QY 586 PAPIEKTISKAGGAPREPQVYVILPSPSEEMIKNOVSLTCLVKGFYPSDIAVEWESNGQPE 645
 DB 583 PAPIEKTISKAGGAPREPQVYVILPSPSEELIKN-VSLTCLVKGFYPSDIAVEWESNGQPE 642
 QY 646 NNYKIIIPVLISSDGSFPIYSKLIIVKSKWQGVNVSQSVMHLEAHNHYIUKSLSISPKR 704
 DB 643 NNYKTIIPVLISSDGSFPIYSKLIIVKSKWQGVNVSQSVMHLEAHNHYIUKSLSISPKR 701

RESULT 3

ID Q95M34 PRELIMINARY; PRT: 337 AA.
 AC Q95M34;
 DT 01-DEC-2001 (Tremblrel, 19, created)
 DT 01-DEC-2001 (Tremblrel, 19, last sequence update)
 DT 01-DEC-2001 (Tremblrel, 19, last annotation update)
 DE IMMUNOGLOBULIN GAMMA 1 HEAVY CHAIN CONSTANT REGION
 DE (FRAGMENT).
 GN IGHCL.
 OS Equus caballus (Horse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
 OX NCBI_TaxID=9796;
 RN 11
 RP SEQUENCE FROM N.A.
 RA Wagner B.;


```

Db 109 R-NTSGKRIIVYKMLQEAALFPELSLITIVNOGENINISYSRLYSPELTVIHKNGHPHS 167
QY 165 VPHFEVPTGLEVILLPDAAPQI-AWYSAPYTGQNLFTSAFTPLIVPPEACQKWPPEPHIL 223
Db 168 SPKEISDTHIVPTNAKESARIVAIYISAAAPSSAAITPLIVRSCPGAFWPNFTES 227
QY 224 CTAAMNNVTHVPTGICITGCTEMPCITKAPITHTFCITCKPSCGCP-CKSYVHCLP 282
Db 228 CFCANGWVGETTLEWVITGCTGHIILAVAGEFCAAPKEP-VIGWVALVCLP 285
QY 283 DEYGSICATGWGLQENFAHNEVGPILKLPSTNNEMDPGLGLSPHWWGLQEP 342
Db 286 DEYGSICATGWGLQENFAHNEVGPILKLPSTNNEMDPGLGLSPHWWGLQEP 344
QY 343 EGIPMTKIVDLPDIIEVNSG-KENPICASGWPLTNEEMTLVKPDGTVLHPKDFNHT 401
Db 344 AD---SSPVLSILRD-VEINTGVLSVNSGASGPPALHQPDTLITANPTTTAAVTHTL 399
QY 402 DHFSVAITPTTHLPPDSGVVWCVSVNTVAGWVEKPNISVKVLKPKLNAIPNVIDTGHNFA 461
Db 406 NGQSTSVKVVQVRV-SAPRWPVQNNTHIMAVHEFTVEKVPVPPQNPPEVLQSGPPHL 454
QY 462 VINISSEPYFGE 473
Db 460 LILLINTEPYSGD 471

RESULT 5
QY9LC4
ID Q99LC4 PRELIMINARY: PRG: 463 AA.
AC Q99LC4:
DT 01-JUN-2001 (Tremblrel, 17, Created)
DT 01-JUN-2001 (Tremblrel, 17, Last sequence update)
DT 01-DEC-2001 (Tremblrel, 19, Last annotation update)
DE SIMILAR TO P1KEN CNA 181066600 GENE
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RC Submitted (FBI-2001) to the EMBL/GenBank/DBJ databases
DR EMBL: AC003435; AA003435.1; -
DR HSSP: P01842; 7FAB.
DR InterPro: IPR003599; Iq.
DR InterPro: IPR003600; Iq_c1.
DR InterPro: IPR003606; Iq_like.
DR InterPro: IPR003006; Iq_MHC.
DR Pfam: PF00047; Iq; 3.
DR SMART: SM00409; Iq; 2.
DR SMART: SM00407; Iq; 2.
DR SMART: SM00406; Iq; 1.
DR SMART: SM00410; Iq_like; 1
DR PROSITE: PS00290; Iq_MHC; UNKNOWN_1.
SQ SEQUENCE 463 AA: 51007 MW: 54864.666030783 CRC64:

Query Match 21.4%; Score 849; DH 11; Length 463;
Best Local Similarity 45.4%; Pred. No. 1,26,69;
Matches 160; Conservative 55; Mismatches 65; Indels 44; Gaps 6;

```

```

QY 415 LPPDSGVWCVSVNTVAGWVEKPNISVKVLKPKL-----NAPNVDTGHNF----- 460
Db 150 LAPGSAATNSMTVIGCLVKQYF-----PEPVITVWNSGSLSSGVHTPAVLQSDLYT 202
QY 461 --AVINISSEPYFGEPKSKDTHI-----CPP--CPAPELLGQPSVFLP 500
Db 203 LSSSVIVPSSTWSPSEIVTCNVAHIPASSTKVKDKIVPDDGCKPCICTVPEV---SSVFIF 259
QY 501 PPKPKDVLIMSKITPVIICVWVWVSHEDPEKFNWYVAGVGVHNAKINPREQVNSYTRVV 560

```

```

Db 260 PPKPKDVLITLTPKVTCCVVDISKDDPEVOF-SNPFVDDVEVHTAQTPREFQNSIFRSV 319
QY 561 SVLTVLHGLWLNSEYKCKVSNKALPAPIETITSKAKGQAPPEFQVYTLPPSPEEMTKNOV 620
Db 320 SELPTMHQDLWLNSEYKCKVSNKALPAPIETITSKAKGQAPPEFQVYTLPPPEKQMAKDKV 379
QY 621 SLTCLVKGFYFSDTAVENSNQGFENNYKTTIPVLDSPGSEELYSKYKIVK-SRWQGNVNF 680
Db 480 SLTCLVKGFYFSDTAVENSNQGFENNYKTTIPVLDSPGSEELYSKYKIVK-SRWQGNVNF 680
QY 681 SCSTMHEALHNHYTOKSLSLSPGK 704
Db 440 TCSVLHGLHNNHTKSLSHSPGK 463

RESULT 6
QY9LIA4
ID Q99LIA4 PRELIMINARY: PRG: 437 AA.
AC Q99LIA4:
DT 01-MAY-2000 (Tremblrel, 13, Created)
DT 01-MAY-2000 (Tremblrel, 13, Last sequence update)
DT 01-DEC-2001 (Tremblrel, 19, Last annotation update)
DE GAMMAL HEAVY CHAIN OF MAB7 (FRAGMENT)
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Wilde K.G.; Yu X.; Kramoddoullah A.K.M.; Misra S.;
RC Cloning of cDNAs encoding for anti-white pine blister rust monoclonal
RC antibody (Mab 7, its light and heavy chains) and construction of a
RC single chain antibody (scFv).
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF152372; AA040243.1; -
DR HSSP: P01842; 7FAB.
DR InterPro: IPR003600; Iq_like.
DR InterPro: IPR003006; Iq_MHC.
DR InterPro: IPR003596; Iq_v.
DR Pfam: PF00047; Iq; 4.
DR SMART: SM00406; Iq; 1.
DR SMART: SM00410; Iq_like; 2.
DR PROSITE: PS00290; Iq_MHC; UNKNOWN_1.
FT NON_TER 1
FT NON_TER 437
SQ SEQUENCE 437 AA: 48142 MW: 50347.833706470 CRC64:

```

```

Query Match 21.8%; Score 845; DH 11; Length 437;
Best Local Similarity 49.1%; Pred. No. 2,56,69;
Matches 159; Conservative 56; Mismatches 65; Indels 44; Gaps 6;
QY 415 LPPDSGVWCVSVNTVAGWVEKPNISVKVLKPKL-----NAPNVDTGHNF----- 460
Db 134 LAPGSAATNSMTVIGCLVKQYF-----PEPVITVWNSGSLSSGVHTPAVLQSDLYT 176
QY 461 --AVINISSEPYFGEPKSKDTHI-----CPP--CPAPELLGQPSVFLP 500
Db 177 LSSSVIVPSSTWSPSEIVTCNVAHIPASSTKVKDKIVPDDGCKPCICTVPEV---SSVFIF 243
QY 501 PPKPKDVLIMSKITPVIICVWVWVSHEDPEKFNWYVAGVGVHNAKINPREQVNSYTRVV 560
Db 234 PPKPKDVLITLTPKVTCCVVDISKDDPEVOF-SNPFVDDVEVHTAQTPREFQNSIFRSV 294
QY 561 SVLTVLHGLWLNSEYKCKVSNKALPAPIETITSKAKGQAPPEFQVYTLPPSPEEMTKNOV 620
Db 294 SKLPTMHQDLWLNSEYKCKVSNKALPAPIETITSKAKGQAPPEFQVYTLPPPEKQMAKDKV 354
QY 621 SLTCLVKGFYFSDTAVENSNQGFENNYKTTIPVLDSPGSEELYSKYKIVK-SRWQGNVNF 680
Db 344 SLTCLVKGFYFSDTAVENSNQGFENNYKTTIPVLDSPGSEELYSKYKIVK-SRWQGNVNF 680
QY 681 SCSTMHEALHNHYTOKSLSLSPGK 704

```


Db 426 FSAVGEASICEEDWNSGERETCTVTHIDLPSPKQJLSEPKVALLHPQVLLPPAPAEQL 486
 QY 616 T-KNWSLITLVKGVPSGIAVWESNGQP--ENNYKTIIPVLD--SIGSPFLYSKLTVLD 670
 Db 486 NUKESATITCLTVGSPALVFWVMQRCQPLSPKQVTSAPMPQAPQACRYFAHSILTVS 545
 QY 671 KSRWQGNVFSQVMHEALHNIYTKSLSPGK 704
 Db 546 EREWNIGETCTVCAVAHEALPNHVTIKRVDKSGIK 579
 RESULT 14
 Q96Y00 PRELIMINARY: PRT: 597 AA.
 ID Q96Y00
 AC Q96Y00
 DT 01-JUN-2001 (TREMURel. 17, Created)
 DT 01-JUN-2001 (TREMURel. 17, Last sequence update)
 DT 01-DEC-2001 (TREMURel. 19, Last annotation update)
 DE HYPOTHETICAL 65.3 KDA PROTEIN.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCRI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LYMPH, LYMPHOMA;
 RA Strausberg R.;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.
 DR EMBL: AC002963; AA002963.1; -;
 DR BSSP: P01825; 7EAB.
 DR InterPro: IPR003594; Iq.
 DR InterPro: IPR003597; Iq_c1.
 DR InterPro: IPR003600; Iq_Like.
 DR InterPro: IPR003006; Iq_MHC.
 DR InterPro: IPRJ03596; Iq_V.
 DR Pfam: PF00047; Iq; 5.
 DR SMART: SM00403; Iq; 2.
 DR SMART: SM00407; Iq; 4.
 DR SMART: SM00408; Iq; 1.
 DR SMART: SM00410; Iq_Like; 1.
 DR PROSITE: PS00290; Iq_MHC; UNKNOWN_3.
 KW Hypothetical protein.
 SQ SEQUENCE 597 AA: 65.34 MW: 21464677.85561 CP:64;

Query Match 9.3%; Score 359; DB 4; Length 597;
 Best Local Similarity 34.1%; Pred. No. 30-24;
 Matches 73; Conservative 48; Mismatches 85; Indels 8; Gaps 6;

QY 497 VELFPPKPKDTLMISRTPEVTCVVVDVSHEDPEVKFNWYVDGVEVHNAKTKPREEQYNST 556
 Db 368 VFAIDPVS-FASITLTKSTKICLTVLDITYD--SVTISWTRONGRAVKTHTNISLSPHNAT 425
 QY 557 YRVVSVLTLLHGDWLNKKEYKGVSKALPAPIEKTIKSKAKGQP-REPOVYTLPPSPREEM 615
 Db 426 ESAVGEASICEEDWNSGERETCTVTHIDLPSPKQJLSEPKVALLHPQVLLPPAPAEQL 485
 QY 616 T-KNWSLITLVKGVPSGIAVWESNGQP--ENNYKTIIPVLD--SIGSPFLYSKLTVLD 670
 Db 486 NUKESATITCLTVGSPALVFWVMQRCQPLSPKQVTSAPMPQAPQACRYFAHSILTVS 545
 QY 671 KSRWQGNVFSQVMHEALHNIYTKSLSPGK 704
 Db 546 EREWNIGETCTVCAVAHEALPNHVTIKRVDKSGIK 579

RESULT 15
 Q96Y00 PRELIMINARY: PRT: 613 AA.
 ID Q96Y00
 AC Q96Y00
 DT 01-DEC-2001 (TREMURel. 19, Created)
 DT 01-DEC-2001 (TREMURel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMURel. 19, Last annotation update)

DE UNKNOWN (PROTEIN FOR M30:20337).
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCRI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=PRIMARY B-CELLS FROM IONISILS;
 RA Strausberg R.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
 DR EMBL: AC011857; AA011857.1; -;
 SQ SEQUENCE 613 AA: 67.23 MW: 31214204.068421E7 CP:64;

Query Match 9.2%; Score 355; DB 4; Length 614;
 Best Local Similarity 21.6%; Pred. No. 7.20-24;
 Matches 156; Conservative 92; Mismatches 249; Indels 246; Gaps 28;
 QY 34 LVSDAET-SLTGCIASWPPHEPITITRFEALMNOHQPPELVTCQVREMAKKVWKEEK 92
 Db 30 LVKPSHLSLICIVSG-----QSISVY-----WS--WIQOP 59
 QY 93 ASKINGAYFCGPGVPCFPAIPPTMPPOQASHPDAPLTMIVKQKNNVNSPKKVLKEED 152
 Db 60 AGK--GLEWIG-----RIYTSSTNPSLSKSRVTSVDTSKN-QFSLKLSSTAAD 108
 QY 153 AVIYKNGSPHSDVRHEVDHLEVHPHQAQDQACVARSYIG-CNIFUSAFHLIVRRP 211
 Db 109 TAVYYCAS-----QWELPTVGLFYWGQGLIVVS-----Y 148
 QY 212 EAGWVPECNILCTACMNGVHEETGECIQFPMETPEKAELETITGRTYERESNQ 271
 Db 139 -----SOSASAPFLPLVSCF-----NNSDSEVVA 164
 QY 272 PGCKSVVFCIPVYCCSCATQCKGLQCNF-ACHQPF-----Y 307
 Db 165 VGLAQDF-LPD-----STFHWYKNNSSISGTFPSVIGKRYAATSQVLLPSKVMQ 219
 QY 308 GPD-----CKLPGSCNNGMCHDFQCCGCPQWQGLQCRREGIPRMPTIKIVLPHIEV 361
 Db 220 GTDHEHWCKVQHPGCKE-----KNVP--LIVLAFLPDKVSFV 256
 QY 362 NSGKFNDFCKASG WFLTNEEMTLVKPDGTVLPIKPDENITDHF 404
 Db 257 PRDGFPGNPKSKILICVATGFSFQIQVSWLRKQKQVSGVITDQVQAKAKSGPTTYK 416
 QY 405 SVAIEITHRLPDPISGVVQSVNTVAGNVEKPIETSVKVLKPLNAPNVIDIGHNFAVIN 464
 Db 317 VISTLTIKESDWLSQSMETCRDII-----RGLTFQONASSCMCPDQDTAI-- 361
 QY 465 ISSEYPFGPKSCNTHICPCPAPELIGGPSVHLFPPKPKDTLMISRTPEVTCVVVWS 524
 Db 362 -----RVFALDPS-FASITLTKSTKICLTVLDIT 389
 QY 525 HEDPEKFNWYVDGVEVHNAKTKPREEQYNSTYRVVSVLTLLHGDWLNKKEYKGVSKNA 584
 Db 390 TYD-SVTISWTRONGRAVKTHTNISLSPHNATPSAVGHASICHEDWNSGRPTCTVTHTD 448
 QY 585 LPAPIEKTIKSKAKGQP-REPOVYTLPPSPREEMT-KNWSLITLVKGVPSDIAVWESNG 642
 Db 449 LPSPLKQTIISRPKGVALHRPDVYLLPPAREQLNRESATITCLTVGSPALVFWVMQRC 508
 QY 643 QP--ENNYKTIIPVLD--SIGSPFLYSKLTVKSRWQGNVFSQVMHEALHNIYTKSL 698
 Db 509 QPLSPKQVTSAPMPQAPQACRYFAHSILTVSEFWNTIGETCTVCAVAHEALPNHVTIKR 548
 QY 699 SLS 701
 Db 569 DKS 571

Search completed: September 23, 2002, 22:12:59

Job time: 195 sec

100

101

102

103

104

105

106

107

108

109

110

111

112

113

114

115

116

117

118

119

120

121

122

123

124

125

126

127

128

129

130

131

132

133

134

135

136

137

138

139

140

141

142

143

144

145

146

147

148

149

150

151

152

153

154

155

156

157

158

159

160

161

162

163

164

165

166

167

168

169

170

171

172

173

174

175

176

177

178

179

180

181

182

183

184

185

186

187

188

189

190

191

192

193

194

195

196

197

198

199

200

201

202

203

204

205

206

207

208

209

210

211

212

213

214

215

216

217

218

219

220

221

222

223

224

225

226

227

228

229

230

231

232

233

234

235

236

237

238

239

240

241

242

243

244

245

246

247

248

249

250

251

252

253

254

255

256

257

258

259

260

261

262

263

264

265

266

267

268

269

270

271

272

273

274

275

276

277

278

279

280

281

282

283

284

285

286

287

288

289

290

291

292

293

294

295

296

297

298

299

300

301

302

303

304

305

306

307

308

309

310

311

312

313

314

315

316

317

318

319

320

321

322

323

324

325

326

327

328

329

330

331

332

333

334

335

336

337

338

339

340

341

342

343

344

345

346

347

348

349

350

351

352

353

354

355

356

357

358

359

360

361

362

363

364

365

366

367

368

369

370

371

372

373

374

375

376

377

378

379

380

381

382

383

384

385

386

387

388

389

390

391

392

393

394

395

396

397

398

399

400

401

402

403

404

405

406

407

408

409

410

411

412

413

414

415

416

417

418

419

420

421

422

423

424

425

426

427

428

429

430

431

432

433

434

435

436

437

438

439

440

441

442

443

444

445

446

447

448

449

450

451

452

453

454

455

456

457

458

459

460

461

462

463

464

465

466

467

468

469

470

471

472

473

474

475

476

477

478

479

480

481

482

483

484

485

486

487

488

489

490

491

492

493

494

495

496

497

498

499

500

501

502

503

504

505

506

507

508

509

510

511

512

513

514

515

516

517

518

519

520

521

522

523

524

525

526

527

528

529

530

531

532

533

534

535

536

537

538

539

540

541

542

543

544

545

546

547

548

549

550

551

552

553

554

555

556

557

558

559

560

561

562

563

564

565

566

567

568

569

570

571

572

573

574

575

576

577

578

579

580

581

582

583

584

585

586

587

588

589

590

591

592

593

594

595

596

597

598

599

600

601

602

603

604

605

606

607

608

609

610

611

612

613

614

615

616

617

618

619

620

621

622

623

624

625

626

627

628

629

630

631

632

633

634

635

636

637

638

639

640

641

642

643

644

645

646

647

648

649

650

651

652

653

654

655

656

657

658

659

660

661

662

663

664

665

666

667

668

669

670

671

672

673

674

675

676

677

678

679

680

681

682

683

684

685

686

687

688

689

690

691

692

693

694

695

696

697

698

699

700

701

702

703

704

705

706

707

708

709

710

711

712

713

714

715

716

717

718

719

720

721

722

723

724

725

726

727

728

729

730

731

732

733

734

735

736

737

738

739

740

741

742

743

744

745

746

747

748

749

750

751

752

753

754

755

756

757

758

759

760

761

762

763

764

765

766

767

768

769

770

771

772

773

774

775

776

777

778

779

780

781

782

783

784

785

786

787

788

789

790

791

792

793

794

795

796

797

798

799

800

801

802

803

804

805

806

807

808

809

810

811

812

813

814

815

816

817

818

819

820

821

822

823

824

825

826

827

828

829

830

831

832

833

834

835

836

837

838

839

840

841

842

843

844

845

846

847

848

849

850

851

852

853

854

855

856

857

858

859

860

861

862

863

864

865

866

867

868

869

870

871

872

873

874

875

876

877

878

879

880

881

882

883

884

885

886

887

888

889

890

891

892

893

894

895

896

897

898

899

900

901

902

903

904

905

906

907

908

909

910

911

912

913

914

915

916

917

918

919

920

921

922

923

924

925

926

927

928

929

930

931

932

933

934

935

936

937

938

939

940

941

942

943

944

945

946

947

948

949

950

951

952

953

954

955

956

957

958

959

960

961

962

963

964

965

966

967

968

969

970

971

972

973

974

975

976

977

978

979

980

981

982

983

984

985

986

987

988

989

990

991

992

993

994

995

996

997

998

999

1000

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

QM protein - protein search, using sw model

Pub. on: September 23, 2002, 11:29 Search time: 00:04 seconds
(without alignments)
1699 962 Million cell updates/sec

Title: US-09-733-764-2_copy_1_472
Perfect score: 2605
Sequence: 1 MDSLASLVGVSLISGV VITSHNPNVINISSEYPPG 472

Scoring table:
Gapop 10.0 Gapext 0.5

Searched: 747574 seqs, 111073796 residues
Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:			
1:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:	*	
2:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:	*	
3:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:	*	
4:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:	*	
5:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:	*	
6:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:	*	
7:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:	*	
8:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:	*	
9:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:	*	
10:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:	*	
11:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:	*	
12:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:	*	
13:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:	*	
14:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:	*	
15:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:	*	
16:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:	*	
17:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:	*	
18:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:	*	
19:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:	*	
20:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:	*	
21:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:	*	
22:	/SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:	*	

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2605	100.0	704	22	AAB31165 Amino acid sequence
2	2605	100.0	977	22	AAB31164 Amino acid sequence
3	2605	100.0	1124	15	AAR45440 Human orphan recep
4	2605	100.0	1124	16	AAR73953 Human TEK tyrosine
5	2605	100.0	1124	20	AA130318 Amino acid sequence
6	2605	100.0	1124	22	AAG65945 Amino acid sequence
7	2407	92.4	1118	21	AA159046 Mouse receptor tyr
8	2407	92.4	1122	21	AA159048 Mouse tek receptor
9	2407	92.4	1123	16	AAR74951 Mouse tie-2 recept
10	2407	92.4	1123	16	AAR73391 Murine tyrosine ki
11	2393	91.9	1122	15	AAR48627 Protein-tyrosine k

12	2391.5	91.8	1117	15	AAR53146	Mouse tyrosine kin
13	2352	90.3	1101	20	AAB31160	Rat orphan tyrosin
14	865.5	33.2	1138	14	AAR39820	tie receptor kinas
15	851	32.7	1137	20	AA130319	Amino acid sequenc
16	844.5	32.4	1122	16	AAR73954	Human tie tyrosine
17	684.5	26.4	736	22	AAG65944	Human human diagno
18	675.5	25.9	1094	14	AAR39821	Truncated tie rece
19	455	27.5	231	20	AA174444	Mouse tie receptor
20	327.5	12.6	434	22	AH866756	Drosophila melanog
21	315	12.1	241	22	AA172712	HTLH44 clone huma
22	315	12.1	384	22	AAG75479	Human colon cancer
23	313.5	12.0	587	22	ABG31826	Novel human diagno
24	313	12.0	878	22	ABG38033	Novel human diagno
25	312.5	12.0	269	22	ABG38031	Novel human diagno
26	307	11.8	497	22	AAB66268	Marine TANGO 272 S
27	307	11.8	436	22	AAR66264	RAF "ANOC 272 SEC
28	305	11.7	474	22	AA172715	HEPC208 clone huma
29	305	11.7	296	22	AAB56069	Skin cell protein,
30	305	11.7	299	21	AA176014	Murine EGF family
31	305	11.7	299	21	AA176077	Murine EGF family
32	305	11.7	299	22	AAB55953	Skin cell protein,
33	305	11.7	299	22	AAB56016	Skin cell protein,
34	300	11.5	310	22	AA172716	HMP1804 clone huma
35	297	11.4	1050	22	AAB66267	Human TANGO 272 SE
36	289.5	11.1	350	22	ABG08075	Novel human diagno
37	286.5	11.0	594	22	AAR58834	Protophila melanog
38	286	11.0	833	21	AA185616	Murine nurse cell
39	286	11.0	833	22	AAG67639	Mouse nurse cell r
40	284.5	10.9	321	22	AAG37639	Novel human diagno
41	277	10.6	129	22	AAG32555	Novel human diagno
42	277	10.6	129	22	ABG27637	Novel human diagno
43	277	10.6	866	22	AAB60393	Human nurse cell r
44	277	10.6	866	22	AAR60394	Human nurse cell r
45	272.5	10.5	466	22	AAG32559	Human human diagno

ALIGNMENTS

RESULT 1					
AAB31165					
1D	AAB31165 standard; protein; 704 AA.				
XX					
AC	AAB31165;				
XX					
DT	02-APR-2001 (first entry)				
XX					
DE	Amino acid sequence of a Tek/Fc fusion protein.				
XX					
KW	Fusion protein; receptor tyrosine kinase; Tek; Fc portion;				
KW	immunoglobulin G; G1, angiogenesis; tumor; cellular neovascularisation;				
KW	inflammatory disease; arthritis; rheumatism; psoriasis; eye disorder;				
KW	diabetic retinopathy; retinopathy; neovascular glaucoma; retinoblastoma;				
KW	retrolental fibroplasia; rubecosis; uveitis; macular degeneration;				
KW	graft neovascularisation; cancer; metastatic sarcoma; carcinoma;				
KW	wound granulation.				
OS	Synthetic.				
OS	Homo sapiens.				
XX					
FT	Key	Location/Qualifiers			
FT	Peptide	1..18			
FT	Protein	/note= "signal peptide"			
FT	Protein	19..472			
FT	Protein	/note= "Tek extracellular domain fragment"			
FT	Protein	473..704			
FT	Protein	/note= "Fc portion"			
XX					
XX	W02000675323-A1.				
XX	14-DEC 2000.				
XX	07-JUN-2000, 2000WO-US15706.				

XX 07-JUN-1999; 990S-0137889.
 XX (IMMV) IMMUNEX CORP.
 XX Carretti DP, Borges LG, Fanslow WC;
 XX WPI; 2001-112149/12.
 XX New Tek polypeptides antagonist having a fragment of the Tek
 PT extracellular domain, useful for treating mammals with a disease
 PT mediated by angiogenesis, e.g. tumors, ocular neovascularisation or
 PT inflammatory diseases -
 XX claim 17; Page 37-39; 43pp; English.
 XX the present sequence represents a fusion protein comprising the
 CC extracellular domain of the human receptor tyrosine kinase Tek and the
 CC Fe portion of human immunoglobulin (Ig) G1. The Tek fragment lacks all
 CC or part of the region containing fibronectin type III (FNIII) motifs, and
 CC retains the ability to bind at least one Tek ligand. The fusion
 CC polypeptide is an angiogenesis inhibitor, and a Tek antagonist. The
 CC polypeptide or soluble Tek multimer, antibodies of antibody fragments are
 CC useful for treating a mammal having a disease or condition mediated by
 CC angiogenesis, e.g. a solid tumor, a condition or disease characterized by
 CC ocular neovascularisation. In particular, the Tek antagonists are useful
 CC for treating or preventing inflammatory diseases (e.g. arthritis,
 CC rheumatism or psoriasis), certain eye disorders (e.g. diabetic
 CC retinopathy, retinopathy of prematurity, neovascular glaucoma,
 CC retinoblastoma, retrolental fibroplasia, rubeosis, uveitis, macular
 CC degeneration or graft neovascularisation), cancer (e.g. metastatic
 CC sarcomas or carcinomas), or wound granulation.
 XX Sequence 704 AA;

Query Match 100.0%; Score 2605; DB 22; Length 704;
 Best Local Similarity 100.0%; Pred. No. 2,2e-187;
 Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDLSASLVCGVSLISGIVEGAMDLIIINSIPILVSDAFTSLITCIASGWRPHEPITICRD 60
 DB 1 mdslaslvicqvslisqvsgamdliinsipilvsdaftslitciassgwrphcpilicrd 60
 QY 61 PPAIMNQHQPPLVTPQVTHWAKKVVWKKPKASKINCAVPCFGRVCEAIPIETMKMQ 120
 DB 61 fealmqhdplievtqdvtrwakkvkvkckaskingayfcoegrvrgeairrtmkmq 120
 QY 121 QASFLPATLTMTVDKIDNVTISFKKVLTKERDAVYKKNSFTHSVPREVPDILEVHLPH 180
 DB 121 qasflpatlmtvdtkydnvtisfkkvltkerdavlykksfthsvprvdpdilevhlph 180
 QY 181 AQWQACVVSARYVCGNIFTSAPTHLIVHPCPAOKWGPCNCHLTACNNNCVCHEDTQPC 240
 DB 181 adpqdaqvsaryvqnltsitstzlvrrceaqkwpcechletacnnngvchedtqec 240
 QY 241 LCPHPCMCRCACAEHIFGCEKSCSCKSVYKCLADAYKCSAAGWKGCTNP 300
 DB 241 lcphtmpcmrcacaehtfgcekscksvykcldadykcsaagwkgctnp 300
 QY 301 ACHPFVYGPTEKLEPSSNNKEMTGFELGTSAPWQSLQAEFEKIPKMTPIVBLPHQIE 360
 DB 301 achpfvygpte klepssnnkemtgfelgtsapwqslqae fekipkmtpi vblphqie 360
 QY 361 VNSCKNPIKCKASQWPIPEFMTIVKQNGIVIHKKPHNHTDHSVAFTIHRILPPNSC 420
 DB 361 vnsqklnpckasqwpipneemtivkpdqvlvhpknhtdhsvalftihrlilppnsq 420
 QY 421 VWVCSVNI VAGWVKEPFIISVAVLPKPLNAPNVIIVTGHNFPAVINISSPPVHG 472
 DB 421 vwvcsvntv agwvkepfisvavlpkplnapnviivtghnfpav inissppvhg 472

RESULT 2
 AAB31164
 ID AAB31164 standard; protein; 977 AA.
 XX
 AC AAB31164;
 XX
 DT 02-APR-2001 (first entry)
 XX
 DE Amino acid sequence of a tek/tye fusion protein.
 XX
 KW Fusion protein, receptor tyrosine kinase; Tek; Fe portion;
 KW immunoglobulin G1; IgG1; angiogenesis; tumor; ocular neovascularisation;
 KW inflammatory disease; arthritis; rheumatism; psoriasis; eye disorder;
 KW diabetic retinopathy; retinopathy; neovascular glaucoma; retinoblastoma;
 KW retrolental fibroplasia; rubeosis; uveitis; macular degeneration;
 KW graft neovascularisation, cancer, metastatic sarcoma, carcinoma;
 KW wound granulation.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 FH Key location/Qualifiers
 FT Peptide 1..18 "signal peptide"
 FT Protein 19..745
 FT Protein /note- "Tek extracellular domain"
 FT Protein 746..977
 FT Protein /note- "Fe portion"
 XX WC200075323-A1.
 XX 14-DEC-2000.
 XX 07-JUN-2000; 2000WO-US15704.
 XX 07-JUN-1999; 990S-0137889.
 XX (IMMV) IMMUNEX CORP.
 XX Carretti DP, Borges LG, Fanslow WC;
 XX WPI; 2001-112149/12.
 XX New Tek polypeptides antagonist having a fragment of the Tek
 PT extracellular domain, useful for treating mammals with a disease
 PT mediated by angiogenesis, e.g. tumors, ocular neovascularisation or
 PT inflammatory diseases -
 XX Claim 41; Page 34-37; 43pp; English.
 XX The present sequence represents a fusion protein comprising the
 CC extracellular domain of the human receptor tyrosine kinase Tek and the
 CC Fe portion of human immunoglobulin (Ig) G1. The Tek fragment lacks all
 CC or part of the region containing fibronectin type III (FNIII) motifs, and
 CC retains the ability to bind at least one Tek ligand. The fusion
 CC polypeptide is an angiogenesis inhibitor, and a Tek antagonist. The
 CC polypeptide or soluble Tek multimer, antibodies or antibody fragments are
 CC useful for treating a mammal having a disease or condition mediated by
 CC angiogenesis, e.g. a solid tumor, a condition or disease characterized by
 CC ocular neovascularisation. In particular, the Tek antagonists are useful
 CC for treating or preventing inflammatory diseases (e.g. arthritis,
 CC rheumatism or psoriasis), certain eye disorders (e.g. diabetic
 CC retinopathy, retinopathy of prematurity, neovascular glaucoma,
 CC retinoblastoma, retrolental fibroplasia, rubeosis, uveitis, macular
 CC degeneration or graft neovascularisation), cancer (e.g. metastatic
 CC sarcomas or carcinomas), or wound granulation.
 XX Sequence 977 AA;

Query Match 100.0%; Score 2605; DB 22; Length 977;
 Best Local Similarity 100.0%; Pred. No. 3,2e-187;
 Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDSLASVLCGVSLLSSTVEGAMDLILNSLPLVSDAETSLTICIASGWKRPHEPITIGRD 60
 DB 1 mdsiasvllcgvslisgsvtegamlilnslplvsdaetslticiasgwrpheapitigrd 60
 QY 61 FEALMNHQDPLEVTDVTPWEAKKVVWKEKASKINGAYFCEGVPGEFAIPTPKMKMQ 120
 DB 61 fealmnhqdplevtdvtpewakkkvwwkekaskingayfcegvpgefaiprtmkmq 120
 QY 121 QASFLPATLMTVKGDNVNISFKKKVILKEEDAVIYKNGSFTHSVPRHEVPDLLEVLPH 180
 DB 121 qasflpatlmtvkgdnvnisfkkvilkedaviykngsfthsvprhevpdllevlph 180
 QY 181 AQPDGAVYSARYIGNLTSAFTPLTVPRPCEAKWGPENHLCCTACMNGVCHEDTGE 240
 DB 181 aqpdgavysaryignltsaftpltvprpceakwgpennhlcctacmngvchedtge 240
 QY 241 ICPPGPMGTCCKACELHTFGTCKPCKSCQDQKSVVPLPVPYQCSAIGCKGLQCNF 300
 DB 241 icppgpmgtckackelhtfgtckpckscdqkssvplpvpqcsaigckglqcnf 300
 QY 301 ACHPGFYGPCKLKCSCNNGEMCDKPGCCICSPQWQGLQCFRKSLVHMTPKIVDLPHIE 360
 DB 301 achpgfygpcklkcscnngemcdkpgccicspqwqglqcfrkslvhmtpkivdlphie 360
 QY 361 VNSGKFNPKCKASGWPLPTNEEMTLVKGDTVLHPKDFNHTDHSVAIFTIHRLLPPDSG 420
 DB 361 vnsqkfnpkckasgwplptneemtlvkgdvtlhpkdfnhtdhsvaiftihrlppdsq 420
 QY 421 VWVCSVNTVAGMVEKPFNISVKVLKPPLNAPVNTDGHNFVAVINISSEPYFG 472
 DB 421 vwvcsvntvagmvekpfnisvkvlpplnapvntdghnfavavinissepfig 472

RESULT 3
 AAR45440
 ID AAR45440 standard; protein: 1124 AA.
 XX AC AAR45440;
 XX DT 25-JUN-1994 (first entry)
 XX DE Human orphan receptor kinase.
 XX KW Ork; ligands; antibodies; PCR; amplification.
 XX OS Homo sapiens.
 XX FI Key Location/Qualifiers
 FT Peptide 1..18
 FT Region /note= "signal peptide"
 FT Region 746..772
 FT Region /note= "transmembrane region"
 FT Region 211..340
 FT Region /note= "ECF-like repeat"
 FT Misc-difference 44
 FT /note= "Cys of immunoglobulin domain"
 FT Misc-difference 102
 FT /note= "Cys of immunoglobulin domain"
 FT Misc-difference 19..1124
 FT /note= "claimed fragment"
 FT Misc-difference 19..745
 FT /note= "claimed fragment"
 XX WO9400469-A.
 XX PD 06-JAN-1994.
 XX PF 25-JUN-1993; 94W0-0506093
 XX PR 26-JUN-1992; 92US 0905600.
 XX PA (IMMUNEX) IMMUNEX COPP

XX Ziegler SE;
 XX WPI: 1994-026132/03.
 XX N-PSDB; AAO55179.
 XX DNA and protein sequences for orphan receptor tyrosine kinase -
 XX and expression vectors for production of recombinant protein and
 XX antibodies specific for the protein, useful in research
 XX Claim 13; Fig 1; 57pp; English.
 XX Degenerate oligonucleotide primers based on the sequence conserved
 XX in the kinase domain of all receptor tyrosine kinases was used for
 XX PCR of single stranded cDNA from human placental polyA mRNA. PCR
 XX prod. HPK-6 contained a novel sequence which was used as a probe to
 XX isolate longer fragments from a human placental cDNA library. One
 XX clone contained the entire coding region and was called the ork gene.
 XX The gene prod. shown can be used as a research tool in in vitro assays
 XX for detection of ork, its ligands or their interactions.
 XX See also AAR45441.
 XX SQ Sequence 1124 AA;
 Query Match 100.0%; Score 2605; DB 15; Length 1124;
 Best Local Similarity 100.0%; Pred. No. 3,8e-187;
 Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MDSLASVLCGVSLLSSTVEGAMDLILNSLPLVSDAETSLTICIASGWKRPHEPITIGRD 60
 DB 1 mdsiasvllcgvslisgsvtegamlilnslplvsdaetslticiasgwrpheapitigrd 60
 QY 61 FEALMNHQDPLEVTDVTPWEAKKVVWKEKASKINGAYFCEGVPGEFAIPTPKMKMQ 120
 DB 61 fealmnhqdplevtdvtpewakkkvwwkekaskingayfcegvpgefaiprtmkmq 120
 QY 121 QASFLPATLMTVKGDNVNISFKKKVILKEEDAVIYKNGSFTHSVPRHEVPDLLEVLPH 180
 DB 121 qasflpatlmtvkgdnvnisfkkvilkedaviykngsfthsvprhevpdllevlph 180
 QY 181 AQPDGAVYSARYIGNLTSAFTPLTVPRPCEAKWGPENHLCCTACMNGVCHEDTGE 240
 DB 181 aqpdgavysaryignltsaftpltvprpceakwgpennhlcctacmngvchedtge 240
 QY 241 ICPPGPMGTCCKACELHTFGTCKPCKSCQDQKSVVPLPVPYQCSAIGCKGLQCNF 300
 DB 241 icppgpmgtckackelhtfgtckpckscdqkssvplpvpqcsaigckglqcnf 300
 QY 301 ACHPGFYGPCKLKCSCNNGEMCDKPGCCICSPQWQGLQCFRKSLVHMTPKIVDLPHIE 360
 DB 301 achpgfygpcklkcscnngemcdkpgccicspqwqglqcfrkslvhmtpkivdlphie 360
 QY 361 VNSGKFNPKCKASGWPLPTNEEMTLVKGDTVLHPKDFNHTDHSVAIFTIHRLLPPDSG 420
 DB 361 vnsqkfnpkckasgwplptneemtlvkgdvtlhpkdfnhtdhsvaiftihrlppdsq 420
 QY 421 VWVCSVNTVAGMVEKPFNISVKVLKPPLNAPVNTDGHNFVAVINISSEPYFG 472
 DB 421 vwvcsvntvagmvekpfnisvkvlpplnapvntdghnfavavinissepfig 472

RESULT 4
 AAR73953
 ID AAR73953 standard; Protein: 1124 AA.
 XX AC AAR73953;
 XX DT 22-JAN-1996 (first entry)
 XX DE Human HK tyrosine kinase protein.
 XX KW tie-2; receptor-tyrosine kinase; DNA primer; cancer; angiogenesis;

KW vasculogenesis: tek.

OS Homo sapiens.

PN WC09513387-A1.

XX 18-MAY-1995.

XX 12-NOV-1994; 94WO-EU03767.

XX 12-NOV-1993; 93US-0152552.

PA (PLAC) MAX PLANCK GES FORDERUNG WISSENSCHAFTEN.

PI Ritsau W;

DR WP1; 1995-194105/25.

XX New tie-2 receptor tyrosine kinase and related nucleic acid and
PT methods for detecting tie-2 modulators for treating eq cancer,
PI associated with angiogenesis and vasculogenesis

XX Disclousure; Page 49; 81pp; English.

XX This protein is the human homolog of mouse tie-2 receptor
CC tyrosine kinase.

XX Sequence 1124 AA;

Query Match 100.0%; Score 2605; DB 16; Length 1124;

Best Local Similarity 100.0%; Prod. No. 3.8e-187;

Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0,

QY 1 MDSLASIVCGVSLISGIVGAMDLILINSPLVSDAETSLTICIASQWRPHEPTIGRO 60

Db 1 mdsaslvcgysllisgtvegamdlilinsplvsdaetslticlasqwrpheapitard 60

QY 61 FTALMNGHDPLEVYQDVTIRWAKKVKASKINCAYPECEKVKCEAIRITKMKRQ 120

Db 61 ftalnmghdpplvqdvtrwakkvkwaskindayfcedvtrgcairirtmkmrq 120

QY 121 QASFLPATITMTVDKGNVNISEPKVLIKEEDAVIYKNGSPFHSVPRHEVPIILEVILPH 180

Db 121 qasflpatirtvdykgnvnisfkkvlikeedavlykngsfhsvprhpvdpillevhlph 180

QY 181 AQPDAGVYSARYIGNIFTSAPTRILVRRCPAQRWKKRQKINCAIMNNGVCHEDTQEC 240

Db 181 aqpdaqvysaryiqnifltsaltclivrrceaqkwqecnhlclacmnngqvchedtqec 240

QY 241 ICNPGMGRNCHKACHETFTGRTCKRQSGQPGCKSVYFCLLIPYQSCAIGWKGICQNF 300

Db 241 icpplmqrckachetftgtrckrksqpgcksvyfccllipyqscatgkwkgicqnf 300

QY 301 ACHPGFYGPQCKLRSPNNGEMETFGTQTSAPQWNRNLPFEFSPMTKIVLSPHITF 360

Db 301 achpgfygpdcklrespnngemetfgtqtsapqwgngnlpfeefspmtkivlsphtf 360

QY 361 VASQKSNFICRNASQWRPINEHMLIVKRGIVLRKPKFNHILATACMNNVAVFIHKRLLPVSQ 420

Db 361 vasqkfnlckasqwrplnehmliivrkgivlvrkpkfnhiltatcmnnvavfihkrllpvsq 420

QY 421 VVWGSVNTVACQVKEFENIEVKVJLKKIINAVLRLCHNFAVNISSLPVFC 472

Db 421 vvwgsvntvqgvkeefenievkvjllkkinavrlrchnfavnisslpvfc 472

RESULT 5

AAV30318

XX AAV30318 standard; protein; 1124 AA.

XX AAV30318;

XX

DI 15-NOV-1999 (first entry)

XX Amino acid sequence of TEK (also known as TIE2) protein.

XX TEK protein; TIE2 protein; receptor tyrosine kinase; T cell response;

KW immune response; endothelial cell; tumor-associated vasculature;

KW coagulation; thrombosis; cancer; anticancer vaccine.

XX Homo sapiens.

XX WO9943801-A1.

XX 02-SEP-1999.

XX 26-FEB-1999; 99WO-CB00583.

XX 26-FEB-1998; 98CH-0004121.

XX (CANC-) CANCER RES CAMPAIGN TECHNOLOGY.

XX Barrant LG, Hewett PW, Ramage JM, Spendlove I;

XX WP1; 1999-540586/45.

XX New peptides containing at least one epitope from Tek receptor
PT tyrosine kinase, used in vaccines against cancer

XX Disclousure; Fig 1; 56pp; English.

XX The present sequence represents the TEK protein, also known as TIE2
CC protein. TEK is a receptor tyrosine kinase. TEK contains epitopes
CC which bind to MHC. The presentation of TEK epitopes can also stimulate
CC helper cell and/or cytotoxic T cell responses. The immune response is
CC directed against endothelial cells in the tumor-associated
CC vasculature and includes production of antibodies that bind to the
CC cells, causing coagulation and thrombosis. The immune response is
CC targeted to endothelial cells lining blood vessels of the tumor (these
CC cells overexpress Tek). So damage to even a few cells will kill many
CC tumor cells. These target cells are accessible to the immune response
CC and problems of antigenic heterogeneity, MHC loss and resistance to
CC apoptosis (associated with epithelial cells) are unlikely to occur in
CC normal endothelial cells. TEK epitopes (see AAY30320-24) are used to
CC generate antibodies and for prevention and treatment of cancer.
CC The peptides, and recombinant DNA constructs or viral vectors that
CC express them, are useful as anticancer vaccines to target endothelial
CC cells that line blood vessels of the tumor.

XX Sequence 1124 AA;

Query Match

Best Local Similarity 100.0%; Score 2605; DB 20; Length 1124;

Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDSLASIVGVSLISGIVGAMDLILINSPLVSDAETSLTICIASQWRPHEPTIGRO 60

Db 1 mdsaslvcgysllisgtvegamdlilinsplvsdaetslticlasqwrpheapitard 60

QY 61 FTALMNGHDPLEVYQDVTIRWAKKVKASKINCAYPECEKVKCEAIRITKMKRQ 120

Db 61 ftalnmghdpplvqdvtrwakkvkwaskindayfcedvtrgcairirtmkmrq 120

QY 121 QASFLPATITMTVDKGNVNISEPKVLIKEEDAVIYKNGSPFHSVPRHEVPIILEVILPH 180

Db 121 qasflpatirtvdykgnvnisfkkvlikeedavlykngsfhsvprhpvdpillevhlph 180

QY 181 AQPDAGVYSARYIGNIFTSAPTRILVRRCPAQRWKKRQKINCAIMNNGVCHEDTQEC 240

Db 181 aqpdaqvysaryiqnifltsaltclivrrceaqkwqecnhlclacmnngqvchedtqec 240

QY 241 ICNPGMGRNCHKACHETFTGRTCKRQSGQPGCKSVYFCLLIPYQSCAIGWKGICQNF 300

Db 241 icpplmqrckachetftgtrckrksqpgcksvyfccllipyqscatgkwkgicqnf 300

QY 301 ACHPFTYSPGCKKASNNEMTLPFPASPTSPSWGLQEPESIPMTPIKIVDLPDITE 460
 DB 301 achpfyfpdcklrescngmedrfgclcspsgwgglqeragipmtpkivdlpdhie 360

QY 361 VNSCKENPICKASQWPLTNEMTLVKPDCTVHPRDFNHTDHFSAITTHRIIPPPNS 420
 DB 361 vnscknppickasqwppltnemtlvkdpcdtvhprdfnhtdhfsaitthrilpppsq 420

QY 421 VVCSVNTVAGMVKPKNISVKVLPKPLNAPNVIDTCHNFAVINISSEYVFG 472
 DB 421 vvcsvntvagmvekpnisvkvlpkplnapnvidtchnfavinissepyfq 472

RESULT 6
 AA65945
 ID AAG65945 standard; protein: 1124 AA.
 AC AAG65945;
 XX
 DT 11-FEB-2002 (first entry)
 DE Amino acid sequence of human Tie-2.
 XX
 KW Tie-2, catalytic domain, receptor tyrosine kinase, crystalline; human;
 KW cytosolic; vasotropic; anti-anemic; antiarteriosclerotic; nephrotropic;
 KW ophthalmologic; hepatotropic; antithyroid; antiinflammatory; antiocler;
 KW gastrointestinal; antirheumatic; osteopathic; antiarthritic; hemostatic;
 KW antipsoriatic; dermatologic; immunosuppressive; antibacterial
 OS Homo sapiens
 XX
 FH Key Location/Qualifiers
 FT Domain 802..1124
 FT /note="catalytic domain, specifically claimed fragment"
 XX
 PN WC200172778-A2.
 PD 04-OCT-2001.
 XX
 PF 20 MAR 2001, 2001WO 0538833.
 XX
 PP 29-MAR 2000; 2000US-192920P
 PA (BAU) BASF AG.
 XX
 PI Rump NJ, Arnold LD, Dixon RW, Hoeffken HW, Ailen K, Bellamacina C;
 DR WPI; 2001-648437/74.
 XX
 PT Crystalline polypeptide useful for identifying inhibitors of a Tie-2
 PT protein as well as determining the three dimensional structure of a
 PT polypeptide comprising the catalytic domain of a Tie-2 polypeptide -
 XX
 PS Claim 6; Fig 1; 242pp; English.
 CC The invention relates to a crystalline polypeptide, comprising the
 CC catalytic domain of a receptor tyrosine kinase Tie-2 protein. The
 CC crystalline forms are useful for identifying inhibitors of a Tie-2
 CC protein as well as determining the three dimensional structure of a
 CC the catalytic domain of a Tie-2 polypeptide. A Tie-2 inhibitor may be
 CC used to treat a Tie-2 dependent condition in a patient (especially a
 CC human), where the condition is characterized by excessive vascular
 CC proliferation e.g. a hyperproliferative disorder, cancer (e.g. sarcoma,
 CC osteoma, melanoma, lymphoma, and leukemia), a cardiovascular condition
 CC (e.g. atherosclerosis, ischemia, anemia, and vascular leakage disorders),
 CC an ocular condition (myopia, chronic retinal detachment, conjunctivitis,
 CC retinopathy, and macular degeneration), von Hippel Lindau disease,
 CC pemphigoid, psoriasis, Paget's disease, polycystic kidney disease,
 CC fibrosis, sarcoidosis, cirrhosis, thyroiditis, Osler-Weber-Rendu disease,
 CC chronic inflammation, synovitis, inflammatory bowel disease, Crohn's
 CC disease, rheumatoid arthritis, osteoarthritis, psoriatic arthritis, and
 CC ulcer or sepsis, especially where the disorder involves aberrant

CC endothelial-periendothelial interactions. The Tie-2 inhibitor may be
 CC used to decrease fertility, and promote angiogenesis or vasculogenesis
 CC (in combination with a pro-angiogenic growth factor). The present
 CC sequence represents the human tie 2 protein sequence.
 XX
 SQ Sequence 1124 AA;
 Query Match 100.0%; Score 2605; DB 22; Length 1124;
 Best local Similarity 100.0%, Pred. No. 3,8e-187;
 Matches 472, Conservative 0, Mismatches 0, Indels 0, Gaps 0;
 QY 1 MDSLASLVGVSVLLSSTVEGAMDILINSLPLVSDAETSLTFCASQWPPPIPTITQPD 60
 DB 1 mdslaslvlgvslisgstvegamdilinslplvsdaetsltfcasgwrpfpitigrd 60
 QY 61 FEALMNOHQHPIFVTVTPFWAKKVVWKKFKASKINGAYFCRGPVPGAIPIPTMKMPC 120
 DB 61 fealmnqhdpdvltqvtrwakkvvwkkaskingayfcoqrvrgeairfntkmq 120
 QY 121 QASFLPATLIMTVDKGNVNISPKKVLIKKEDAVIYKNGSFTHSVPRHVPDILVHLPH 180
 DB 121 qasflpatlmtvdkgnvnistkklkcedavlykngsfthsvprhvpdilevhlph 180
 QY 141 ALPLIAVYSAFYTNELFSATPLIVPRKEAKWQENGLLIAWNSVYHEDTCEP 240
 DB 141 alpqdagvysaryggalfisatrlivrtcoagkwgpcchictacmngvchedtgee 240
 QY 241 LPPQPMARCEKACELHIGRICRERCSQVFCQSVVPLKPVVYGCSCATQWKGQCNE 300
 DB 241 lppqpmarceakcelhigrickercsqeqaksyvfclpdygscatqwkqlqene 300
 QY 361 AHPGFTYSPGCKKASNNEMTLPFPASPTSPSWGLQEPESIPMTPIKIVDLPDITE 360
 DB 361 achpfyfpdcklrescngmedrfgclcspsgwgglqeragipmtpkivdlpdhie 360
 QY 361 VNSCKENPICKASQWPLTNEMTLVKPDCTVHPRDFNHTDHFSAITTHRIIPPPNS 420
 DB 361 vnscknppickasqwppltnemtlvkdpcdtvhprdfnhtdhfsaitthrilpppsq 420
 QY 421 VVCSVNTVAGMVKPKNISVKVLPKPLNAPNVIDTCHNFAVINISSEYVFG 472
 DB 421 vvcsvntvagmvekpnisvkvlpkplnapnvidtchnfavinissepyfq 472

RESULT 7
 AA59046
 ID AA59046 standard; protein: 1118 AA.
 AC AA59046;
 XX
 DT 07-MAR-2000 (first entry)
 DE Mouse receptor tyrosine kinase protein.
 XX
 KW Receptor tyrosine kinase protein; angiogenesis; cardiogenesis; mouse;
 KW tumorigenesis.
 XX
 OS Mus musculus.
 XX
 PN M55998187-A
 XX
 PD 07-DEC-1999.
 XX
 PF 23-APR-1997; 97US-0838957.
 XX
 PP 20-MIL-1404; 94NS-0278089.
 PP 40-ILL-1992; 92NS-0021795.
 PP 29-APR-1994; 94US-0235408.
 XX
 PA (MOUN) MOUNT SINAI HOSPITAL CORP.
 XX
 PI Yamauchi TP, Breitman T, Dumont DT, Passant J, Breitman ML;

Best Local Similarity 91.9%, Score 2344, PE 15; Length 1122;
Matches 432; Conservative 17; Mismatches 23; Indels 0; Gaps 0;

QY 1 MDSIASLVLCVSVLLSCTVFCAMDLILNSIPVLSVAEISLTCTASGWRPHPHITIGRD 60
DB 1 MDSIAQLVLCQSVLLYQVVEQAMDLILNSIPVLSVAEISLTCTASGWRPHPHITIGRD 60
QY 61 FEALMNHQDPLEVTQVTPPEWAKKVVWKKPKASKINCAVFCGKRVCEAIPITMKMPQ 120
DB 61 FEALMNHQDPLEVTQVTPPEWAKKVVWKKPKASKINCAVFCGKRVCEAIPITMKMPQ 120
QY 121 QASFLPATLTMVTDGDNVNISFKVLKKEEDAVIYKNGSEIHSVPRIEVPDILEVHLPH 180
DB 121 QASFLPATLTMVTDGDNVNISFKVLKKEEDAVIYKNGSEIHSVPRIEVPDILEVHLPH 180
QY 181 AQPDAGVYSARYIGCNLFTSAFTPLVPRCEAKWGPCNHLCTACMNNVCHEHVGEC 240
DB 181 AQPDAGVYSARYIGCNLFTSAFTPLVPRCEAKWGPCNHLCTACMNNVCHEHVGEC 240
QY 241 ICPFGFMGRTECEKAGELHIFGRICKERGSGGEGKSVYFGLPDYGGSCAIGWKLQUNE 300
DB 241 ICPFGFMGRTECEKAGELHIFGRICKERGSGGEGKSVYFGLPDYGGSCAIGWKLQUNE 300
QY 301 ACPHGFYGDCKLRSCNNGEWTFFPGGTCSPGWGLQVPEEIPMTPTKPIVDLPHIE 360
DB 301 ACPHGFYGDCKLRSCNNGEWTFFPGGTCSPGWGLQVPEEIPMTPTKPIVDLPHIE 360
QY 361 VNSCKFNPICKASGWPIDTNEMTLVKPDGTVLHPKPNHRTDHSVAIPTIHRITLPHDSG 420
DB 361 VNSCKFNPICKASGWPIDTNEMTLVKPDGTVLHPKPNHRTDHSVAIPTIHRITLPHDSG 420
QY 421 VVWCSVNTVAGMVEKPPNISVKKVLPKPLNAPNVIDTGHNFAVINISSEPYFG 472
DB 421 VVWCSVNTVAGMVEKPPNISVKKVLPKPLNAPNVIDTGHNFAVINISSEPYFG 472

RESULT 11
AAR48627
ID AAR48627 standard; Protein; 1122 AA.
AC AAR48627;
XX AAR48627;
DT 31-JUL-1994 (first entry)
DE protein-tyrosine-kinase tek.
XX protein-tyrosine kinase, tek transcriptional regulatory element,
KW endothelium; gene therapy; vascular disease.
XX Mus musculus.
XX W09404694-A.
XX 03-MAR-1994.
XX 25-AUG-1993; 93WO-CA00352.
XX 25-AUG-1992; 92US 0944393
XX (MOON) MOUNT SINAI HOSPITAL CORP.
XX Breitman ML, Dumont D, Gradwohl GG;
XX WPI; 1994-083208/10.
XX N-PSDB; AAO56697.
PT Isolated transcriptional regulatory element - for directing
PT expression of gene specifically in cells of endothelial lineage
XX
XX Disclosure; Page 42-46; 74pp; English.
XX cDNA of sequence AAO56697, derived from mouse strain CD 1 embryo
XX heart, encodes a protein-tyrosine-kinase (tek) of sequence AAR48627

CC that is expressed during mouse cardiogenesis. An additional cDNA
CC sequence, given in AAO56697 and encoding protein AAR48628, was isolated
CC from a phage lambda-gt10 mouse embryo cDNA library. DNA of sequence
CC AAO56697 was isolated from a mouse genomic phage library using tek
CC cDNA as probe, and included the transcriptional regulatory element
CC of tek comprising the initiation codon and untranslated sequences.
CC The element may be used in gene therapy to introduce foreign genes
CC into endothelial cells to correct/prevent vascular disease.
XX
SQ Sequence 1122 AA;

Query Match 91.9%, Score 2344, PE 15; Length 1122;
Best Local Similarity 91.5%, Pred No 3 2e-171;
Matches 432; Conservative 17; Mismatches 23; Indels 0; Gaps 0;

QY 1 MDSIASLVLCVSVLLSCTVFCAMDLILNSIPVLSVAEISLTCTASGWRPHPHITIGRD 60
DB 1 MDSIAQLVLCQSVLLYQVVEQAMDLILNSIPVLSVAEISLTCTASGWRPHPHITIGRD 60
QY 61 FEALMNHQDPLEVTQVTPPEWAKKVVWKKPKASKINCAVFCGKRVCEAIPITMKMPQ 120
DB 61 FEALMNHQDPLEVTQVTPPEWAKKVVWKKPKASKINCAVFCGKRVCEAIPITMKMPQ 120
QY 121 QASFLPATLTMVTDGDNVNISFKVLKKEEDAVIYKNGSEIHSVPRIEVPDILEVHLPH 180
DB 121 QASFLPATLTMVTDGDNVNISFKVLKKEEDAVIYKNGSEIHSVPRIEVPDILEVHLPH 180
QY 181 AQPDAGVYSARYIGCNLFTSAFTPLVPRCEAKWGPCNHLCTACMNNVCHEHVGEC 240
DB 181 AQPDAGVYSARYIGCNLFTSAFTPLVPRCEAKWGPCNHLCTACMNNVCHEHVGEC 240
QY 241 ICPFGFMGRTECEKAGELHIFGRICKERGSGGEGKSVYFGLPDYGGSCAIGWKLQUNE 300
DB 241 ICPFGFMGRTECEKAGELHIFGRICKERGSGGEGKSVYFGLPDYGGSCAIGWKLQUNE 300
QY 301 ACPHGFYGDCKLRSCNNGEWTFFPGGTCSPGWGLQVPEEIPMTPTKPIVDLPHIE 360
DB 301 ACPHGFYGDCKLRSCNNGEWTFFPGGTCSPGWGLQVPEEIPMTPTKPIVDLPHIE 360
QY 361 VNSCKFNPICKASGWPIDTNEMTLVKPDGTVLHPKPNHRTDHSVAIPTIHRITLPHDSG 420
DB 361 VNSCKFNPICKASGWPIDTNEMTLVKPDGTVLHPKPNHRTDHSVAIPTIHRITLPHDSG 420
QY 421 VVWCSVNTVAGMVEKPPNISVKKVLPKPLNAPNVIDTGHNFAVINISSEPYFG 472
DB 421 VVWCSVNTVAGMVEKPPNISVKKVLPKPLNAPNVIDTGHNFAVINISSEPYFG 472

RESULT 12
AAR53146
ID AAR53146 standard; Protein; 1117 AA.
XX AAR53146;
XX AAR53146;
DT 21-NOV-1994 (first entry)
XX Mouse tyrosine kinase ("tek") expressed during cardiogenesis.
XX Endothelial specific receptor tyrosine kinase; PTF; tek gene;
KW mouse; murine; chromosome 4; cardiogenesis.
XX
XX Mus musculus.
XX
XX Key location/Qualifiers
XX Domain 821..1117
XX /label- catalytic_domain
XX Region 821..839
XX /label- conserved_region_1
XX /note- "the tek amino acid sequence appears
XX elsewhere in the specification with
XX the residues FQDV between positions
XX 822 and 823"

FT Region 840...857
 FT /label= conserved_region_II
 FT 873...929
 FT /note= "GXGXXG motif,
 FT part of consensus ATP-binding site"
 FT Binding-site 846...848
 FT /note= "AXK motif,
 FT part of consensus ATP-binding site"
 FT Region 858...873
 FT /label= conserved_region_III
 FT Region 874...888
 FT /label= conserved_region_IV
 FT Region 889...907
 FT /label= conserved_region_V
 FT Region 908...928
 FT /label= insert
 FT Region 929...969
 FT /label= conserved_region_VI
 FT Region 970...988
 FT /label= conserved_region_VII
 FT Region 975...977
 FT /note= "motif found in almost all known kinases"
 FT Region 989...1003
 FT /label= conserved_region_VIII
 FT Region 997...1003
 FT /note= "motif found in transmembrane RTKs"
 FT Region 1004...1034
 FT /label= conserved_region_IX
 FT Region 1035...1055
 FT /label= conserved_region_X
 FT /note= "the tek amino acid sequence appears
 FT e somewhere in the specification with
 FT the residues Et at positions 1039-1040
 FT rather than DV"
 FT Region 1056...1085
 FT /label= conserved_region_XI
 FT Region 1086...1117
 FT /label= conserved_region_XII
 XX CAZ085291-A.
 XX 31-JAN-1994.
 XX 14-DEC-1992; 92CA-2085291.
 XX 30-JUL-1992; 92US-0921795.
 XX (MOUN) MOUNT SINA; HOSPITAL CORP.
 XX Breitman ML, Dumont DJ, Rossant J, Yamaguchi TP;
 XX WPI: 1994-126938/16.
 XX N-PSDB: AAQ62129.
 PT Receptor tyrosine kinase isolated from cells of the endothelial
 PT lineage - used for identifying (ant)agonists of ligand-receptor
 PT binding
 XX Claim 9: Fig 1: 51pp; English.
 XX RNA from 9.5 and 12.5 day old mice was used to synthesize cDNA by RT
 CC PCR using primers previously demonstrated to preferentially amplify
 CC tyrosine kinase sequences. Clones corresp. to 5 distinct tyrosine
 CC kinases expressed during murine cardiogenesis were identified. Four
 CC were known tyrosine kinases but one represented a new tyrosine kinase
 CC which was designated "tek". AAK53146 is the deduced tek amino acid
 CC sequence. Comparisons to other tyrosine kinases showed that the
 CC putative kinase domain contains several conserved sequence motifs
 CC (see features table).
 XX Sequence 1117 AA:

Query Match 91.8%; Score 2391.5; DR 15; Length 1117;
 Best local Similarity 91.7%; Pred. No. 4, 10-171;
 Matches 433; Conserved 17; Mismatches 21; Indels 1; Gaps 1;
 QY 1 MDSLASLVLCVSVLLSGTVEGAMDLILNSPLVSDAETSLTCTASGWPPEPTIGRD 60
 Db 1 mdsilagivlcgvslllygvvgamdliinsplvsdaetsltctasgwppeptigrd 60
 QY 61 FEALMNGHQPPLFVTVQVTPFWAKKVVWPKFASKINKAYFGGSPVGFPAIPIPTMKWQ 120
 Db 61 fealmnqhqpplfvvtqvtrwakkvwwkrekaskinqayicegrvrgqairlrmkqrq 120
 QY 121 QASPLPATLMTVSKGNVN:SPKKVLIRKEDAVIKNSPFIHSVPRHVPDILVHLPH 180
 Db 121 qastlpaltlmtvdrqdnvnlstkkvlikeedaviknsfihsvprhvpdilevhlph 180
 QY 181 ACPALAVYSAFYLG:NLFTSAFTPI:VPPPEALPWTFE:NIILCLIA:MNNVW:HEUT:EC 240
 Db 181 acpalavysafylg:nlftsaftpi:vpppealpwtfeniilclia:mnnvw:heut:ec 240
 QY 241 ICPG:FWK:PI:EKAL:ELHTR:KPK:KES:SQ:PS:KSVPR:LP:PY:AS:AT:WKL:QV:NE 400
 Db 241 icpg:fwk:pi:ekal:elhtr:kpk:kessq:ps:ksvpr:lp:py:as:at:wkl:qv:ne 400
 QY 301 ACHG:PYG:PUCKLKCSCNNGEMCDK:QGCLCS:PWG:GLQ:CK:KG:IP:KMI:PK:IV:DL:PH:LE 360
 Db 300 acpsqyygppcklrchctneelcdriqgclcsqgqlqckeqgrprmpqiedipdhle 359
 QY 361 VNSGFKNPICASQWPIPTNFEMTLVKPQGTVLIRPKDFNHTDHF:SV:AI:FT:HP:TI:PP:SG 420
 Db 360 vnsqkfnpickasqwpptseentlvkpqutvlirpndfntdfrfsvaiftvnrviipdsq 419
 QY 421 VWVQSVNTVAGVMEKPFNISVKVLPRPI:NPAPNVITGHNF:AV:IN:SS:EP:YFG 472
 Db 420 vwvcsvatvagmekpfnisvk:lpeptkapnvitghnfaivinssepyfg 471
 RESULT 13
 AAW83160
 ID AAW83160 standard; Protein; 1101 AA.
 XX
 AC AAW83160;
 XX
 DT 11-FEB-1999 (first entry)
 XX
 DE Rat orphan tyrosine kinase receptor protein Rtk 7 (Tie 2).
 XX Receptor tyrosine kinase, Ror 1, Ror 2, Etk-1; Etk-2, detection,
 KW neurotrophin activity; trkB; proto-oncogene; tyrosine kinase receptor;
 KW binding protein; bDNF; Nf-3; diagnosis.
 XX
 OS Rattus sp.
 XX
 PN US5843749-A.
 XX
 PD 01-DEC-1998.
 XX
 PF 06-JUN-1995; 95US-0469537.
 XX
 PR 17-MAR-1995; 95US-0406247.
 PF 26-JUL-1991; 91US-0736559.
 PR 28-OCT-1993; 93US-0144992.
 PF 06-JUN-1995; 95US-0469537.
 XX
 PA (PENSE-) REGENERON PHARM INC
 XX
 PI Maisongier PC, Masiakowski P, Yancopoulos GB;
 XX WPI: 1999-044584/04.
 XX
 PT DNA encoding receptor tyrosine kinase proteins - and corresponding
 PT proteins
 XX

```

PS Example: Fig 20; 194pp; Enlish.
XX
CC The present invention describes nucleic acid molecules for ror-1,
CC ror-2, chk 1 and chk 2. Also described are the corresponding proteins;
CC ror-1, ror-2, chk 1, and chk 2. The proteins are orphan tyrosine
CC tyrosine kinases. The present sequence represents rat orphan tyrosine
CC kinase receptor Rtk-7 (Tie-2) from the present invention.
XX
SQ Sequence 110; AA:

Query Match 90.3%; Score 2352; DB 20; Length 1101;
Best Local Similarity 93.4%; Pred No 3 8e-168;
Matches 423; Conservative 12; Mismatches 18; Indels 5; Gaps 0;

QY 20 VEGAMDLILINSPLVSVDAETSLTCLASQWRHPHPTITIGRDFEALMNOHQDPLEVTDVVT 79
Db 1 veyamdliinsplvsdaetstlclsqwphphptitigrdfealmnohqdpplvtqdv 60
QY 80 REWAKKVVWKREKASKINAYFEGVGRGEATPTITMKRQCAFLPATLTMTYDKGDNV 139
Db 1 rewakkvwwrkaskingayfegvgrgeatptitmkrcqasflpatltmtvdygdav 120
QY 140 NISFKKVLKEEDAVIYKNGSFHISVPREVPDILEVILPHAQPDAGVYSARYIGGNLF 199
Db 1 nisfkvllkeedaviyknasfhisvprevpdilevilphaqpdagvysaryiggnlf 180
QY 200 ISAFTRILVRRCEAKWSPCEACNAIACMNNQVCHEDIQEGICHPGFMGRICENACEIHT 259
Db 1 isatfrilvrceakwspceacnaicmnnqvcchediqegichpgfmgriccekacephl 240
QY 260 FHPYTKFPGSQEGVKSVEFTPTFYSVSSATLAKWGLQENAYHPSVYSPMLKLVSCNN 319
Db 1 fhpkykfgsqegvksvftptfysvssatlwkwglqenayhpsvyspmlklvscnn 300
QY 241 ftrtkeressgsgsksyvflpdyssasatlwglqenayhpsvyspmlklvscnn 300
QY 320 GEMDFQFGLCSQWGLQAFEPETIPMTIKIVLPBHEVNSKFNPICKASGWPILT 379
Db 1 gemdfqfglcsqwglqafepetipmtikivlpbhevnsgkfnpicckasgwplpt 360
QY 301 eemcdfqlqcsqwpqlqekarpmtqjedlphievsqklupickasgwplpt 360
QY 380 NEMTLVKVIGTIVLHKIPNHTDHSVAFTIHKILPVDSCVWVCSVNTVAGVVEKPFNI 439
Db 1 nemtlvkvigtivilhkippnhtdhsvaftihkilpvdscvwwvcsvntvagvvekpfin 420
QY 440 SVKVLKPKLPANVITDITGNEAVINISSEPYFG 472
Db 1 svkvlkpklpnvnitdittgneaavinissepyfg 453

RESULT 14
AAR39820
XX AAR39820 standard; Protein: 1138 AA.
XX
AC AAR39820;
XX
DT 12-JAN-1994 (first entry)
XX
DE tie receptor kinase.
XX
KW Tie; receptor, tyrosine kinase, endothelial cell; immunoglobulin; lig-
KW epidermal growth factor; EGF; repeat; blood vessel; atherosclerosis;
KW thromboembolic disease; neoplastic disease; tumour angiogenesis;
KW wound healing; inflammatory disease.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..21
FT Protein /note= "Signal peptide"
FT Protein 22..1138
FT Modified-site 83..85 /note= "Mature protein"
FT Modified-site 161..163 /note= "N-linked glycosylation site"

```

```

FT Domain /note= "N-linked glycosylation site"
FT 214..257
FT FT /note= "EGFH-I"
FT 258..304
FT 305..346
FT Domain /note= "EGFH-II"
FT 305..346
FT FT /note= "EGFH-III"
FT 503..505
FT Modified-site /note= "N-linked glycosylation site"
FT 503..505
FT FT /note= "N-linked glycosylation site"
FT 596..598
FT Modified-site /note= "N-linked glycosylation site"
FT 709..711
FT FT /note= "N-linked glycosylation site"
FT 761..786
FT Region /note= "Transmembrane region"
FT 837..1107
FT Domain /note= "Tyrosine kinase domain"
FT 938..951
FT FT /note= "Kinase insert"
FT 938..951
XX W09314124-A.
PN 22-JUL-1993.
XX
XX 08-JAN-1993; G4W0-FI00006.
XX
XX 09-JAN-1992; 92US-0817R00.
XX (UYHF-) UNIV HELSINKI HOLDING LTD.
XX
XX Alltalo K, Armstrong E, Korhonen J, Makela TP;
XX Partanen J;
XX WPI; 1993-243152/30.
XX N-PSDB; AA047064.
XX
XX Nucleic acid encoding tyrosine kinase receptor - for regulating
XX tie activity in diagnosis and treatment of neoplastic diseases
XX involving tumour angiogenesis, wound healing etc.
XX
XX Claim 2; Page 40-44; 73pp; English.
XX
XX The sequence given in AAR39820 represents the full length 'tie' receptor
XX kinase. The sequence given in AAR39821 represents a truncated
XX version of this, wherein residues corresponding to positions 214
XX 257 of the first sequence are absent. Tie is an endothelial cell
XX receptor tyrosine kinase and contains immunoglobulin (Ig-) and
XX epidermal growth factor (EGF-) like repeats. The tie gene is
XX expressed in the endothelial cells of blood vessels. Tie levels may
XX indicate onset/progression of diseases involving endothelial cells
XX and their tie receptors, such as neoplastic diseases involving tumour
XX angiogenesis, wound healing, thromboembolic diseases, atherosclerosis
XX and inflammatory diseases.
XX
SQ Sequence 1138 AA:

```

```

Query Match 33.2%; Score 865.5; DB 14; Length 1138;
Best Local Similarity 38.2%; Pred No 1.8e-56;
Matches 174; Conservative 71; Mismatches 185; Indels 25; Gaps 8;

QY 14 LLLSGTVEGAMDLILINSPLVSVDAETSLTCLASQWRHPHPTITIGRDFEAL 64
Db 13 lflashvgaavdltilanrltdpqrfltcvgaagrgsdawqp-plilckdriv 70
QY 65 MNQHQGPVFTVQDTPFWAKVWKREFFASKINAYFEGVGRGEATPTITMKRQCAFLPATLTMTYDKGDNV 124
Db 71 rppppfpr karqshqtlrskpsdlsqvlceagadartrrlychaspah 126
QY 124 FLPATLTMTYDKGDNVNSFKKVLKEEDAVIYKNGSFHISVPREVPDILEVILPH 180
Db 127 flpshvllckadla.lsur.lkekq' dvlwksarsy7yldwheadarflilq lpm 184

```


100

101

102

103

104

105

106

107

108

109

110

111

112

113

114

115

116

117

118

119

120

121

122

123

124

125

126

127

128

129

130

131

132

133

134

135

136

137

138

139

140

141

142

143

144

145

146

147

148

149

150

151

152

153

154

155

156

157

158

159

160

161

162

163

164

165

166

167

168

169

170

171

172

173

174

175

176

177

178

179

180

181

182

183

184

185

186

187

188

189

190

191

192

193

194

195

196

197

198

199

200

201

202

203

204

205

206

207

208

209

210

211

212

213

214

215

216

217

218

219

220

221

222

223

224

225

226

227

228

229

230

231

232

233

234

235

236

237

238

239

240

241

242

243

244

245

246

247

248

249

250

251

252

253

254

255

256

257

258

259

260

261

262

263

264

265

266

267

268

269

270

271

272

273

274

275

276

277

278

279

280

281

282

283

284

285

286

287

288

289

290

291

292

293

294

295

296

297

298

299

300

301

302

303

304

305

306

307

308

309

310

311

312

313

314

315

316

317

318

319

320

321

322

323

324

325

326

327

328

329

330

331

332

333

334

335

336

337

338

339

340

341

342

343

344

345

346

347

348

349

350

351

352

353

354

355

356

357

358

359

360

361

362

363

364

365

366

367

368

369

370

371

372

373

374

375

376

377

378

379

380

381

382

383

384

385

386

387

388

389

390

391

392

393

394

395

396

397

398

399

400

401

402

403

404

405

406

407

408

409

410

411

412

413

414

415

416

417

418

419

420

421

422

423

424

425

426

427

428

429

430

431

432

433

434

435

436

437

438

439

440

441

442

443

444

445

446

447

448

449

450

451

452

453

454

455

456

457

458

459

460

461

462

463

464

465

466

467

468

469

470

471

472

473

474

475

476

477

478

479

480

481

482

483

484

485

486

487

488

489

490

491

492

493

494

495

496

497

498

499

500

501

502

503

504

505

506

507

508

509

510

511

512

513

514

515

516

517

518

519

520

521

522

523

524

525

526

527

528

529

530

531

532

533

534

535

536

537

538

539

540

541

542

543

544

545

546

547

548

549

550

551

552

553

554

555

556

557

558

559

560

561

562

563

564

565

566

567

568

569

570

571

572

573

574

575

576

577

578

579

580

581

582

583

584

585

586

587

588

589

590

591

592

593

594

595

596

597

598

599

600

601

602

603

604

605

606

607

608

609

610

611

612

613

614

615

616

617

618

619

620

621

622

623

624

625

626

627

628

629

630

631

632

633

634

635

636

637

638

639

640

641

642

643

644

645

646

647

648

649

650

651

652

653

654

655

656

657

658

659

660

661

662

663

664

665

666

667

668

669

670

671

672

673

674

675

676

677

678

679

680

681

682

683

684

685

686

687

688

689

690

691

692

693

694

695

696

697

698

699

700

701

702

703

704

705

706

707

708

709

710

711

712

713

714

715

716

717

718

719

720

721

722

723

724

725

726

727

728

729

730

731

732

733

734

735

736

737

738

739

740

741

742

743

744

745

746

747

748

749

750

751

752

753

754

755

756

757

758

759

760

761

762

763

764

765

766

767

768

769

770

771

772

773

774

775

776

777

778

779

780

781

782

783

784

785

786

787

788

789

790

791

792

793

794

795

796

797

798

799

800

801

802

803

804

805

806

807

808

809

810

811

812

813

814

815

816

817

818

819

820

821

822

823

824

825

826

827

828

829

830

831

832

833

834

835

836

837

838

839

840

841

842

843

844

845

846

847

848

849

850

851

852

853

854

855

856

857

858

859

860

861

862

863

864

865

866

867

868

869

870

871

872

873

874

875

876

877

878

879

880

881

882

883

884

885

886

887

888

889

890

891

892

893

894

895

896

897

898

899

900

901

902

903

904

905

906

907

908

909

910

911

912

913

914

915

916

917

918

919

920

921

922

923

924

925

926

927

928

929

930

931

932

933

934

935

936

937

938

939

940

941

942

943

944

945

946

947

948

949

950

951

952

953

954

955

956

957

958

959

960

961

962

963

964

965

966

967

968

969

970

971

972

973

974

975

976

977

978

979

980

981

982

983

984

985

986

987

988

989

990

991

992

993

994

995

996

997

998

999

1000

Gapcore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 23, 2002, 22:13:40, Search time: 11:42 seconds
(without alignments)
724,176 Million cell updates/sec

Title: US-09-733-764-2_COPY_1_472
Perfect score: 2605
Sequence: 1 MCSLASLVLRGVSLLISCTV VITGRINFAVINTSEPPYPS 472

Scoring table: BLOSUM62
Gapop 10 0, Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Issued Patents.AA.*

- 1: /cqn2_6/ptedata/1/1aa/5A.COMB.pcp.*
- 2: /cqn2_6/ptedata/1/1aa/5B.COMB.pcp.*
- 3: /cqn2_6/ptedata/1/1aa/6A.COMB.pcp.*
- 4: /cqn2_6/ptedata/1/1aa/6B.COMB.pcp.*
- 5: /cqn2_6/ptedata/1/1aa/PCTUS.COMB.pcp.*
- 6: /cqn2_6/ptedata/1/1aa/backfiles1.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2605	100.0	1124	1	US-08-323-474-2
2	2605	100.0	1124	5	PCT US93-06093-2
3	2407	92.4	1118	1	US 07-934-393B-2
4	2407	92.4	1118	1	US-08-278-089A-2
5	2407	92.4	1118	2	US-08-838-957A-2
6	2407	92.4	1122	1	US 08 278-089A-6
7	2407	92.4	1122	2	US-08-838-957A-6
8	2352	90.3	1101	2	US-08-469-537A-96
9	865.5	33.2	1138	1	US-08-323-474-8
10	865.5	33.2	1138	2	US-08-469-537A-98
11	865.5	33.2	1138	2	US 08-220-240A-5
12	814	31.3	1135	2	US-08-469-537A-97
13	455	17.5	131	2	US-08-650-598-3
14	305	11.7	294	4	US-09-188-930-332
15	273	11.4	294	4	US-09 188-930-192
16	266	10.4	47	1	US-08-278-089A-14
17	269	10.3	47	2	US-08-838-957A-18
18	242	9.3	44	1	US-08-278-089A-18
19	242	9.3	44	2	US-08-838-957A-17
20	231.5	8.9	685	3	US-08-872-855-2
21	210	8.1	42	1	US-08-278-089A-20
22	203.5	7.8	2471	1	US 08 185-432-15
23	203.5	7.8	2471	1	US-08-083-590A-19
24	203.5	7.8	2471	3	US-08-532-384-19
25	200.5	7.7	1193	2	US 08-400-159-10
26	200.5	7.7	1193	4	US 08 611-729A-10
27	199.5	7.7	1404	2	US-08-400-159-2

ALIGNMENTS

RESULT 1
US-08-323-474-2
: Sequence 2, Application US/08323474
: Patent No. 5447860
: GENERAL INFORMATION:
: APPLICANT: Ziegler, Steven F.
: TITLE OF INVENTION: NOVEL TYROSINE KINASE
: NUMBER OF SEQUENCES: 8
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Immunex Corporation
: STREET: 51 University Street
: CITY: Seattle
: STATE: Washington
: COUNTRY: US
: ZIP: 98101
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/323,474
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/905,600
: FILING DATE: 26-JUN-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: Seese, Kathryn A.
: REGISTRATION NUMBER: 32,172
: REFERENCE/TITLE NUMBER: 2609
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (206) 587-0430
: TELEFAX: (206) 233-0644
: TLEX: 756822
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1134 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US 08-323 474-2

Query Match 100.0%, Score 2605, DB 1, Length 1124;
Best Local Similarity 100.0%, Pctid No. 1,3e-217;
Matches 472, Conservation 0, Mismatches 0, Gaps 0;
1 MCSLASLVLRGVSLLISCTVITGRINFAVINTSEPPYPS 472
|||||

Db 1 MDSLASLVAGVSLLSSTVTCAMDLILINSLPLVSDAETSLTCTASGWRPHEPITIGRD 60
QY 61 FEALMNHQDPLEVTDVTRWAKKVVWKKRASKINCAVFCGRVCEAIFRTKMKMQ 120
Db 61 FEALMNHQDPLEVTDVTRWAKKVVWKKRASKINCAVFCGRVCEAIFRTKMKMQ 120
QY 121 QASFLPATLTMVDKGDVNIISFKKVLKEEDAVIYKNGSFTHSVPRHEVPDILEVHLPH 180
Db 121 QASFLPATLTMVDKGDVNIISFKKVLKEEDAVIYKNGSFTHSVPRHEVPDILEVHLPH 180
QY 181 AQPDAGVYSARYIGENFTSAFTPLIVPRCEAKWGPENHLTA'MNNVSVHEETGEC 240
Db 181 AQPDAGVYSARYIGENFTSAFTPLIVPRCEAKWGPENHLTA'MNNVSVHEETGEC 240
QY 241 ICPPGFMGRTECKACEHLHTFGRTEKFCRSCGQGGCKSVFCLPDYGGSCALGKGLQCN 300
Db 241 ICPPGFMGRTECKACEHLHTFGRTEKFCRSCGQGGCKSVFCLPDYGGSCALGKGLQCN 300
QY 301 ACHPGFYGHDCKLRCSNNGEMCDRFOGCLGSPQWQGLQCFREGIPRMTPKIVDLPHIE 360
Db 301 ACHPGFYGHDCKLRCSNNGEMCDRFOGCLGSPQWQGLQCFREGIPRMTPKIVDLPHIE 360
QY 361 VNSGKFNPICKASGWLPTNEEMTLVKPDGTVLHPKDFNHTDHSVAIFTHIRLLPPDSG 420
Db 361 VNSGKFNPICKASGWLPTNEEMTLVKPDGTVLHPKDFNHTDHSVAIFTHIRLLPPDSG 420
QY 421 VVVCSTNTVAGVVKRPFNISVKVLPKPLNAPNVIDTGNFAVINISSEPYEG 472
Db 421 VVVCSTNTVAGVVKRPFNISVKVLPKPLNAPNVIDTGNFAVINISSEPYEG 472

RESULT 2

US-07-934-06093-2
Sequence 2, Application PC/US93406093
GENERAL INFORMATION:
APPLICANT: Ziegler, Steven F.
TITLE OF INVENTION: NADEL TYROSINE KINASE
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Immunox Corporation
STREET: 51 University Street
CITY: Seattle
STATE: Washington
COUNTRY: US
ZIP: 98101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCI/US93/06093
FILING DATE: 19940625
CLASSIFICATION:
PRIOR APPLICATION NUMBER: US 07/905,600
FILING DATE: 26-JUN-1992
ATTORNEY/AGENT INFORMATION:
NAME: Seese, Kathryn A.
REGISTRATION NUMBER: 32,172
REFERENCE/DOCKET NUMBER: 2609
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 587-0430
TELEFAX: (206) 233-0644
TELEX: 756422
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1124 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT US93-06093-2

Query Match 100.0%; Score 2605; DB 5; Length 1124;
Rest Local Similarity 100.0%; Pred No. 1.3e 217;
Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDSLASLVAGVSLLSSTVTCAMDLILINSLPLVSDAETSLTCTASGWRPHEPITIGRD 60
Db 1 MDSLASLVAGVSLLSSTVTCAMDLILINSLPLVSDAETSLTCTASGWRPHEPITIGRD 60
QY 61 FEALMNHQDPLEVTDVTRWAKKVVWKKRASKINCAVFCGRVCEAIFRTKMKMQ 120
Db 61 FEALMNHQDPLEVTDVTRWAKKVVWKKRASKINCAVFCGRVCEAIFRTKMKMQ 120
QY 121 QASFLPATLTMVDKGDVNIISFKKVLKEEDAVIYKNGSFTHSVPRHEVPDILEVHLPH 180
Db 121 QASFLPATLTMVDKGDVNIISFKKVLKEEDAVIYKNGSFTHSVPRHEVPDILEVHLPH 180
QY 181 AQPDAGVYSARYIGENFTSAFTPLIVPRCEAKWGPENHLTA'MNNVSVHEETGEC 240
Db 181 AQPDAGVYSARYIGENFTSAFTPLIVPRCEAKWGPENHLTA'MNNVSVHEETGEC 240
QY 241 ICPPGFMGRTECKACEHLHTFGRTEKFCRSCGQGGCKSVFCLPDYGGSCALGKGLQCN 300
Db 241 ICPPGFMGRTECKACEHLHTFGRTEKFCRSCGQGGCKSVFCLPDYGGSCALGKGLQCN 300
QY 301 ACHPGFYGHDCKLRCSNNGEMCDRFOGCLGSPQWQGLQCFREGIPRMTPKIVDLPHIE 360
Db 301 ACHPGFYGHDCKLRCSNNGEMCDRFOGCLGSPQWQGLQCFREGIPRMTPKIVDLPHIE 360
QY 361 VNSGKFNPICKASGWLPTNEEMTLVKPDGTVLHPKDFNHTDHSVAIFTHIRLLPPDSG 420
Db 361 VNSGKFNPICKASGWLPTNEEMTLVKPDGTVLHPKDFNHTDHSVAIFTHIRLLPPDSG 420
QY 421 VVVCSTNTVAGVVKRPFNISVKVLPKPLNAPNVIDTGNFAVINISSEPYEG 472
Db 421 VVVCSTNTVAGVVKRPFNISVKVLPKPLNAPNVIDTGNFAVINISSEPYEG 472

RESULT 3

US-07-934-393H-2
Sequence 2, Application US/07934393H
Patent No. 5466596
GENERAL INFORMATION:
APPLICANT: BREITMAN, MARTIN L.
APPLICANT: DUMONT, DANIEL
APPLICANT: GRADWOHL, GERARD G.
TITLE OF INVENTION: ISSUE SPECIFIC TRANSCRIPTIONAL
ELEMENT
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: HERESKIN & PARK
STREET: 40 King Street West
CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5H 3Y2
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07934,393H
FILING DATE: 25 AUG 1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kurdydk, Linda M.
REGISTRATION NUMBER: 34,971
REFERENCE/DOCKET NUMBER: 3153-64
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 454-7311
TELEFAX: (416) 361-1398
INFORMATION FOR SEQ ID NO: 2:

```

: SEQUENCE CHARACTERISTICS:
: LENGTH: 1118 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-07-934-393B-2

```

```

Query Match: 92.4%; Score 2407; DH 1; length 1118;
Best local Similarity 91.9%; Pred. No. 2c-200;
Matches 434; Conservative 17; Mismatches 21; Indels 0; Gaps 0;

```

```

QY 1 MDSLASLVLCVGLLSSTVSGAMDLILINSIPVSDAETSLTCTIASWHPHEPITIGRD 60
DB 1 MDSLAGLVLCVGLVGVVGGAMDLILINSIPVSDAETSLTCTIASWHPHEPITIGRD 60
QY 61 FEALMNOHQDPLEVTGQVTFEWAQKVVWKKRASKINAYFCEGVVCEGALPIPTMKMPQ 120
DB 61 FEALMNOHQDPLEVTGQVTFEWAQKVVWKKRASKINAYFCEGVVCEGALPIPTMKMPQ 120
QY 121 QASFLPATLTMTVDKGNVNI SFKKVLIKEEDAVIYKNGSFTHSVPRHEVPDILKVLPH 180
DB 121 QASFLPATLTMTVDKGNVNI SFKKVLIKEEDAVIYKNGSFTHSVPRHEVPDILKVLPH 180
QY 181 AQPDAGVSARYIGNLFTSAFTKLIIVPVEAUKWGPENHLCTACMNNVCHEDTGE 240
DB 181 AQPDAGVSARYIGNLFTSAFTKLIIVPVEAUKWGPENHLCTACMNNVCHEDTGE 240
QY 241 ICPPGFMGRTECAKACELHTFGRTCKEPCSGQGGCKSYVFCJLPDYPGCSATGKGLQNE 300
DB 241 ICPPGFMGRTECAKACELHTFGRTCKEPCSGQGGCKSYVFCJLPDYPGCSATGKGLQNE 300
QY 301 ACHPGYGPDPCKLRCSCNNGMCDRPGGCGSGWGGLOCHREGIPRMTPKIVDLPDHIE 360
DB 301 ACHPGYGPDPCKLRCSCNNGMCDRPGGCGSGWGGLOCHREGIPRMTPKIVDLPDHIE 360
QY 361 VNSGKNPICKASGWPDPTEEMTLVKPDGTVLHPKDNHIDHFSVAITFTIHRILPPDSG 420
DB 361 VNSGKNPICKASGWPDPTEEMTLVKPDGTVLHPKDNHIDHFSVAITFTIHRILPPDSG 420
QY 421 VWCVSVNTVAGMVEKPFNISVKVLKPLNAPNVITDGNHFAVINISSEPYFG 472
DB 421 VWCVSVNTVAGMVEKPFNISVKVLKPLNAPNVITDGNHFAVINISSEPYFG 472

```

```

RESULT 4
US-09-733-764-2
: Sequence 2, Application US/09-733-764-2
: Patent No. 5681714
: GENERAL INFORMATION:
: APPLICANT: Breitman, Martin L.
: APPLICANT: Rossant, Janet
: APPLICANT: Dumont, Daniel J.
: APPLICANT: Yamauchi, Terry P.
: TITLE OF INVENTION: No. 5681714e1 Receptor Tyrosine Kinase
: NUMBER OF SEQUENCES: 33
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Hecreskin & Parr
: STREET: 40 King Street West
: CITY: Toronto
: STATE: Ontario
: COUNTRY: Canada
: ZIP: M5H 3Y2
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09-733-764-2
: FILING DATE: 20-JUL-1994
: CLASSIFICATION: 530
: ATTORNEY/AGENT INFORMATION:

```

```

: NAME: Kurdzyk, Linda M.
: REGISTRATION NUMBER: 34,971
: REFERENCE/DOCKET NUMBER: 3153-111
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (416) 364-7311
: TELEFAX: (416) 361-1398

```

```

: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1118 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-09-733-764-2

```

```

Query Match: 92.4%; Score 2407; DH 1; length 1118;
Best local Similarity 91.9%; Pred. No. 2c-200;
Matches 434; Conservative 17; Mismatches 21; Indels 0; Gaps 0;

```

```

QY 1 MDSLASLVLCVGLLSSTVSGAMDLILINSIPVSDAETSLTCTIASWHPHEPITIGRD 60
DB 1 MDSLAGLVLCVGLVGVVGGAMDLILINSIPVSDAETSLTCTIASWHPHEPITIGRD 60
QY 61 FEALMNOHQDPLEVTGQVTFEWAQKVVWKKRASKINAYFCEGVVCEGALPIPTMKMPQ 120
DB 61 FEALMNOHQDPLEVTGQVTFEWAQKVVWKKRASKINAYFCEGVVCEGALPIPTMKMPQ 120
QY 121 QASFLPATLTMTVDKGNVNI SFKKVLIKEEDAVIYKNGSFTHSVPRHEVPDILKVLPH 180
DB 121 QASFLPATLTMTVDKGNVNI SFKKVLIKEEDAVIYKNGSFTHSVPRHEVPDILKVLPH 180
QY 181 AQPDAGVSARYIGNLFTSAFTKLIIVPVEAUKWGPENHLCTACMNNVCHEDTGE 240
DB 181 AQPDAGVSARYIGNLFTSAFTKLIIVPVEAUKWGPENHLCTACMNNVCHEDTGE 240
QY 241 ICPPGFMGRTECAKACELHTFGRTCKEPCSGQGGCKSYVFCJLPDYPGCSATGKGLQNE 300
DB 241 ICPPGFMGRTECAKACELHTFGRTCKEPCSGQGGCKSYVFCJLPDYPGCSATGKGLQNE 300
QY 301 ACHPGYGPDPCKLRCSCNNGMCDRPGGCGSGWGGLOCHREGIPRMTPKIVDLPDHIE 360
DB 301 ACHPGYGPDPCKLRCSCNNGMCDRPGGCGSGWGGLOCHREGIPRMTPKIVDLPDHIE 360
QY 361 VNSGKNPICKASGWPDPTEEMTLVKPDGTVLHPKDNHIDHFSVAITFTIHRILPPDSG 420
DB 361 VNSGKNPICKASGWPDPTEEMTLVKPDGTVLHPKDNHIDHFSVAITFTIHRILPPDSG 420
QY 421 VWCVSVNTVAGMVEKPFNISVKVLKPLNAPNVITDGNHFAVINISSEPYFG 472
DB 421 VWCVSVNTVAGMVEKPFNISVKVLKPLNAPNVITDGNHFAVINISSEPYFG 472

```

```

RESULT 5
US-09-733-764-2
: Sequence 2, Application US/09-733-764-2
: Patent No. 5998187
: GENERAL INFORMATION:
: APPLICANT: Breitman, Martin L.
: APPLICANT: Rossant, Janet
: APPLICANT: Dumont, Daniel J.
: APPLICANT: Yamauchi, Terry P.
: TITLE OF INVENTION: No. 5998187e1 Receptor Tyrosine Kinase
: NUMBER OF SEQUENCES: 32
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Hecreskin & Parr
: STREET: 40 King Street West
: CITY: Toronto
: STATE: Ontario
: COUNTRY: Canada
: ZIP: M5H 3Y2
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible

```

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/838, 957A
FILING DATE: 23-APR-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kurdydyk, Linda M.
REGISTRATION NUMBER: 34,971
REFERENCE/DOCKET NUMBER: 4153-212
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 364-7311
TELEFAX: (416) 361-1398
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1118 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-838-957A-2

Query Match 92.4%, Score 2407, DB 2, length 1118,
Best Local Similarity 91.9%, Fred. No. 2e 200;
Matches 434; Conservative 17; Mismatches 21; Indels 0; Gaps 0;

QY 1 MSLASLVLCVSVLISGTVGAMDLILINSPLVSDAETSLTICASQWPHPEPTIGRD 60
DB 1 MSLAGVLGCVSVLLYGVVGGAMDLILINSPLVSDAETSLTICASQWPHPEPTIGRD 60

QY 61 FEALNNHQDPLEVTDVTPWAKKVVWKKKASKINAYFCEGVPVSEAIRITKMPQ 120
DB 61 FEALNNHQDPLEVTDVTPWAKKVVWKKKASKINCAVEFCQPVVQCOAIRITKMRQ 120

QY 121 QASFLPATLITVDKGNVNISEFKVLLKEEDAVIYKNGSFTHSVPRHPEVDILEVILPH 180
DB 121 QASFLPATLITVDKGNVNISEFKVLLKEEDAVIYKNGSFTHSVPRHPEVDILEVILPH 180

QY 181 AQPDAGVYSARYIGGNLFTSAFTPLIVRRCEAKQWGPCNHLCTACMNNVCVCHETGEC 240
DB 181 AQPDAGVYSARYIGGNLFTSAFTPLIVRRCEAKQWGPCNHLCTACMNNVCVCHETGEC 240

QY 241 ICPPGFMGRTCEKAELELITFRTCKEKSQGGKSKSVFCLPDPYGSATGKGLQCN 300
DB 241 ICPPGFMGRTCEKAELELITFRTCKEKSQGGKSKSVFCLPDPYGSATGKGLQCN 300

QY 301 ACPHGFYGHCKLRCSNNGMCHPEQCCICSPGWGCIQCHREGIPRMTPKIVDLPHIE 360
DB 301 ACPHGFYGHCKLRCHMTNREITGPFQGLTSSQWGLQCEKEGKPPMTPTQIEDLPHIE 360

QY 361 VNSGKFNPKASQWPLINEMILVKDGIVLHPKDFNHDHFSVAIFTIHRILPDSG 420
DB 361 VNSGKFNPKASQWPLISEMTLVKDGIVLHPKDFNHDHFSVAIFTVNRVLPDSG 420

QY 421 VWCVSNTVACWVKPKPNISVKVLPKPLNAPNVDITGHNFAVINISSEPEFG 472
DB 421 VWCVSNTVACWVKPKPNISVKVLPKPLNAPNVDITGHNFAVINISSEPEFG 472

RESULT 6
US-08-838-957A-6
Sequence 6, Application US/08/838, 957A
Patent No. 5681714
GENERAL INFORMATION:
APPLICANT: Breitman, Martin L.
APPLICANT: Rossant, Janet
APPLICANT: Dumont, Daniel J.
APPLICANT: Yamaguchi, Terry P.
TITLE OF INVENTION: No. 5681714cl Receptor Tyrosine Kinase
NUMBER OF SEQUENCES: 33
CORRESPONDENCE ADDRESS:
ADDRESSEE: Bereskin & Parr
STREET: 40 King Street West

CITY: Toronto
STATE: Ontario
COUNTRY: Canada
ZIP: M5H 3Y2
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/278, 089A
FILING DATE: 20-JUL-1994
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Kurdydyk, Linda M.
REGISTRATION NUMBER: 34,971
REFERENCE/DOCKET NUMBER: 3153-111
TELECOMMUNICATION INFORMATION:
TELEPHONE: (416) 364-7311
TELEFAX: (416) 361-1398
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 1122 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-278-089A-6

Query Match 92.4%, Score 2407, DB 1, Length 1122;
Best Local Similarity 91.9%, Fred. No. 2e 200;
Matches 434; Conservative 17; Mismatches 21; Indels 0; Gaps 0;

QY 1 MSLASLVLCVSVLISGTVGAMDLILINSPLVSDAETSLTICASQWPHPEPTIGRD 60
DB 1 MSLAGVLGCVSVLLYGVVGGAMDLILINSPLVSDAETSLTICASQWPHPEPTIGRD 60

QY 61 FEALNNHQDPLEVTDVTPWAKKVVWKKKASKINAYFCEGVPVSEAIRITKMPQ 120
DB 61 FEALNNHQDPLEVTDVTPWAKKVVWKKKASKINCAVEFCQPVVQCOAIRITKMRQ 120

QY 121 QASFLPATLITVDKGNVNISEFKVLLKEEDAVIYKNGSFTHSVPRHPEVDILEVILPH 180
DB 121 QASFLPATLITVDKGNVNISEFKVLLKEEDAVIYKNGSFTHSVPRHPEVDILEVILPH 180

QY 181 AQPDAGVYSARYIGGNLFTSAFTPLIVRRCEAKQWGPCNHLCTACMNNVCVCHETGEC 240
DB 181 AQPDAGVYSARYIGGNLFTSAFTPLIVRRCEAKQWGPCNHLCTACMNNVCVCHETGEC 240

QY 241 ICPPGFMGRTCEKAELELITFRTCKEKSQGGKSKSVFCLPDPYGSATGKGLQCN 300
DB 241 ICPPGFMGRTCEKAELELITFRTCKEKSQGGKSKSVFCLPDPYGSATGKGLQCN 300

QY 301 ACPHGFYGHCKLRCSNNGMCHPEQCCICSPGWGCIQCHREGIPRMTPKIVDLPHIE 360
DB 301 ACPHGFYGHCKLRCHMTNREITGPFQGLTSSQWGLQCEKEGKPPMTPTQIEDLPHIE 360

QY 361 VNSGKFNPKASQWPLINEMILVKDGIVLHPKDFNHDHFSVAIFTIHRILPDSG 420
DB 361 VNSGKFNPKASQWPLISEMTLVKDGIVLHPKDFNHDHFSVAIFTVNRVLPDSG 420

QY 421 VWCVSNTVACWVKPKPNISVKVLPKPLNAPNVDITGHNFAVINISSEPEFG 472
DB 421 VWCVSNTVACWVKPKPNISVKVLPKPLNAPNVDITGHNFAVINISSEPEFG 472

RESULT 7
US-08-838-957A-6
Sequence 6, Application US/08/838, 957A
Patent No. 5998187
GENERAL INFORMATION:
APPLICANT: Breitman, Martin L.
APPLICANT: Rossant, Janet

APPLICANT: Dumont, Daniel J.
 APPLICANT: Yamauchi, Terry P.
 TITLE OF INVENTION: No. 598187el Receptor Tyrosine Kinase
 NUMBER OF SEQUENCES: 32
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Bereskin & Parr
 CITY: Toronto
 STATE: Ontario
 COUNTRY: Canada
 ZIP: M5H 3Y2
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/838,957A
 FILING DATE: 23 APR-1997
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Kutydyk, Linda M.
 REGISTRATION NUMBER: 34,971
 REFERENCE/DOCKET NUMBER: 3153-212
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (416) 364-7311
 TELEFAX: (416) 361-1398
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1122 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-838-957A-6

Query Match 92.4%, Score 2407, DB 2; Length 1122.
 Best Local Similarity 91.9%; Pred. No. 2e-200;
 Matches 434, Conservative 17; Mismatches 21; Indels 0; Gaps 0;

QY 1 MDSIASVLGCVSLLSGTVHGAMDLLINSLPLVSDAETSLTCTASGWRPHEPITIGRD 60
 DB 1 MDSLGLVLCGVSLLGVVEGAMDLLINSLPLVSDAETSLTCTASGWRPHEPITIGRD 60
 QY 61 FEALMNQHDPLVTVQVTRWAKKVVWKKPEKASK INGAYFCEGRVGRGAIRITMKMPQ 120
 DB 61 FEALMNQHDPLVTVQVTRWAKKVVWKKPEKASK INGAYFCEGRVGRGAIRITMKMPQ 120
 QY 121 QASELPATLITVDPKGNVNSFKKVLKKEEDAVIYKNGSFTHSVPRHEVPDILEVHLPH 180
 DB 121 QASELPATLITVDPKGNVNSFKKVLKKEEDAVIYKNGSFTHSVPRHEVPDILEVHLPH 180
 QY 181 ACPTGACVYSAPY IODNIPISATPIIVVPCFACKWGPENHICITACMNKVCHEDTGCE 240
 DB 181 ACPTGACVYSAPY IODNIPISATPIIVVPCFACKWGPENHICITACMNKVCHEDTGCE 240
 QY 241 ICPFGMGRTCEKACELHRTGCKRCSCGGCKSVYFCLDPYGCSCAICWKGLQCNF 300
 DB 241 ICPFGMGRTCEKACELHRTGCKRCSCGGCKSVYFCLDPYGCSCAICWKGLQCNF 300
 QY 301 AHHFEGEYSPK KLKLSNNSEMIDREWGLLSPKAGLGLERESIPPMTPKIVGLPDHLE 360
 DB 301 AHHFEGEYSPK KLKLSNNSEMIDREWGLLSPKAGLGLERESIPPMTPKIVGLPDHLE 360
 QY 361 VNSCKFNPIKASGWPILPFRMTLVKPDCTVILHPKDFNHTDFSAIITTHRLPPDSG 420
 DB 361 VNSCKFNPIKASGWPILPFRMTLVKPDCTVILHPKDFNHTDFSAIITTHRLPPDSG 420
 QY 421 VVCSVNTVAGMVEKPNISVKVLKPLNAPNVNIDTGHNFVAINISSEYFG 472
 DB 421 VVCSVNTVAGMVEKPNISVKVLKPLNAPNVNIDTGHNFVAINISSEYFG 472

RESULT 8
 US-08-469-537A-96
 Sequence 96, Application US/08469537A
 Patent No. 5843749
 GENERAL INFORMATION:
 APPLICANT: Maisongierre, et al.
 TITLE OF INVENTION: EHK AND FOR TYROSINE
 TITLE OF INVENTION: KINASES
 NUMBER OF SEQUENCES: 107
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Regeneron Pharmaceuticals, Inc.
 STREET: 777 Old Saw Mill River Road
 CITY: Tarrytown
 STATE: NY
 COUNTRY: U.S.A.
 ZIP: 10591
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: PASTSQ Version 2.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/469,537A
 FILING DATE: 06-JUN-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: USSN 08/406,247
 FILING DATE: 17-MAR-1995
 APPLICATION NUMBER: USSN 08/144,992
 FILING DATE: 28-OCT-1993
 APPLICATION NUMBER: USSN 07/736,559
 FILING DATE: 26-JUL-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Kempler, Ph.D., Gail M
 REGISTRATION NUMBER: 32,143
 REFERENCE/DOCKET NUMBER: REG 070C
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 914-345-7400
 TELEFAX: 914-345-7721
 TELEX:
 INFORMATION FOR SEQ ID NO: 96:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1101 amino acids
 TYPE: amino acid
 STRANDEDNESS: unknown
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 US-08-469-537A-96

Query Match 90.3%, Score 2352; DB 2; Length 1101;
 Best Local Similarity 91.4%; Pred. No. 1.2e-195;
 Matches 423, Conservative 12; Mismatches 18; Indels 0; Gaps 0;

QY 20 VEGAMDILLINSLPLVSDAETSLTCTASGWRPHEPITIGRDFAALMNQHDPLEVTQDVT 79
 DB 1 VEGAMDILLINSLPLVSDAETSLTCTASGWRPHEPITIGRDFAALMNQHDPLEVTQDVT 60
 QY 80 KEWAKKVVWKKPEKASK INGAYFCEGRVGRGAIRITMKKROOASFLPATLITMTVQKDNV 139
 DB 61 KEWAKKVVWKKPEKASK INGAYFCEGRVGRGAIRITMKKROOASFLPATLITMTVQKDNV 120
 QY 140 NISPKKVLKKEEDAVIYKNGSFTHSVPRHEVPDILEVHLPHAQDPQAGVYSARY IGCNLF 199
 DB 121 NISPKKVLKKEEDAVIYKNGSFTHSVPRHEVPDILEVHLPHAQDPQAGVYSARY IGCNLF 180
 QY 200 TSAETRLIVVPCFACKWGPENHICITACMNKVCHEDTGCEICPPGMORTCEKACEPT 259
 DB 181 TSAETRLIVVPCFACKWGPENHICITACMNKVCHEDTGCEICPPGMORTCEKACEPT 240
 QY 260 FGRICKKRCSCGGCKSVYFCLDPYGCSCAICWKGLQCNFAPYHPCFYSCHNPKIPCSNN 319
 DB 241 FGRICKKRCSCGGCKSVYFCLDPYGCSCAICWKGLQCNFAPYHPCFYSCHNPKIPCSNN 300


```

Query Match      33.2%; Score 865.5; DB 2; Length 1138;
Best Local Similarity 38.2%; Pred. No. 1.2e-66;
Matches 174; Conservative 71; Mismatches 185; Gaps 8;

QY 14 LLLSGTVEGAMDLILLNSLPLVSDAHTSLTCTI-----ASGWRPHPEPTITIGKDFAL 64
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 13 LFLASHVGAADVLTLLANPLTDPPFFITVSGREAGAGPSDAMWCP--PLLEKDDRIV 70
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 65 MNQODPLEVTOVTRFWAKVVKR--EKASKINCAVFCGCRVGRGFAIRIRIMKMQOAS 123
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 71 RTPGPPLR-----LARGSHQVTLRGFSKPSDLGVFSCVGCAGARRTVIVVHNSPGAH 126
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 124 FLPATLMTVDKGNVNISFKKVLKEEDAVIYKNGSFTHSVPRIEVDP---ILEVHLPH 180
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 127 LLEPKVTHVTKGDIATVLSARVHKRQGVVWKSNGSYETILDFAGUGRFLLQ--LPLN 184
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 181 AQPQDACVYSARYICGNLPTSAFTRILIVRCAQAKWGPCNHLCTACMNGVCHEDTQIC 240
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 185 VQPPSSGIYSNIYLEASPLGSAFFELIVGQVNA--PWGQVTFWQVLLQVW--HHDGEC 244
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 241 LCPGPMGRITCEKACELHTFGRICKRCSQDQDQKSYVFLCPYQCCSCATGKMGLOQNE 300
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 245 VCPGFTGTCTGECACNFRGQSCQCCQCPGISGTRGLFCLLPYQCCSGSGSWRGSQOUE 304
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 301 ACHEGFGVSPKLRQSN--NN--EM--TQPPQVTLSPQWQSLQTEPESLPPKMLPKIVLPDHL 360
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 305 ACAPGPHGALVTRIGCCQDQDQDQDQDQDQDQDQDQDQDQDQDQDQDQDQDQDQDQD 361
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 361 VNSGKNPI--CKASGWPLPTNEEMTLVKPDGTVLHPKDFNHTDHFVSVAITFTHRLPPDS 419
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 362 FNLETMPRIN--AAAGNPPFPVPGSIELPKPGDGVLLSTKALIVEPEKTTAEFEVPLVLADS 421
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 420 GVWVCSVNTVAGWVKEKFNISVVLKPKLNAINVI 454
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 422 GPWECRVSTSGQSQSRKFKVNVKVPVPIAAPRL 456
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 11
US-08-220 240A-5
: Sequence 5, Application US/08220240A
: Patent No. 5955291
: GENERAL INFORMATION:
: APPLICANT: Allitalo, Karl
: APPLICANT: Matikainen, Marja-Terttu
: APPLICANT: Partanen, Juhia
: APPLICANT: Makela, Tomi
: APPLICANT: Korhonen, Jaana
: TITLE OF INVENTION: ANTIBIOTICS RESISTINIZING TIE RECEPTOR
: NUMBER OF SEQUENCES: 5
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Horun
: STREET: 213 South Wacker Drive/6400 South Tower
: CITY: Chicago
: STATE: Illinois
: COUNTRY: United States of America
: ZIP: 60606-6402
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM pc compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1 0, Version #1 30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/220, 240A
: FILING DATE: 29-MAR-1994
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PFI/Fiqq/00006
: FILING DATE: 08-JAN-1993
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/817, 800
: FILING DATE: 09-JAN-1992
: PRIOR APPLICATION DATA:

```

```

: APPLICATION NUMBER: US 08/167,453
: FILING DATE: 15-DEC-1993
: ATTORNEY/AGENCY INFORMATION:
: NAME: Cass, David A. 38,153
: REGISTRATION NUMBER: 29151/31968
: REFERENCE/DOCKET NUMBER:
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (312) 474-6300
: TELEFAX: (312) 474-0448
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1138 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE: viral protein
: US-08-220-240A-5

Query Match      33.2%; Score 865.5; DB 2; Length 1138;
Best Local Similarity 38.2%; Pred. No. 1.2e-66;
Matches 174; Conservative 71; Mismatches 185; Gaps 8;

QY 14 LLLSGTVEGAMDLILLNSLPLVSDAHTSLTCTI-----ASGWRPHPEPTITIGKDFAL 64
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 13 LFLASHVGAADVLTLLANPLTDPPFFITVSGREAGAGPSDAMWCP--PLLEKDDRIV 70
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 65 MNQODPLEVTOVTRFWAKVVKR--EKASKINCAVFCGCRVGRGFAIRIRIMKMQOAS 123
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 71 RTPGPPLR-----LARGSHQVTLRGFSKPSDLGVFSCVGCAGARRTVIVVHNSPGAH 126
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 124 FLPATLMTVDKGNVNISFKKVLKEEDAVIYKNGSFTHSVPRIEVDP---ILEVHLPH 180
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 127 LLEPKVTHVTKGDIATVLSARVHKRQGVVWKSNGSYETILDFAGUGRFLLQ--LPLN 184
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 181 AQPQDACVYSARYICGNLPTSAFTRILIVRCAQAKWGPCNHLCTACMNGVCHEDTQIC 240
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 185 VQPPSSGIYSNIYLEASPLGSAFFELIVGQVNA--PWGQVTFWQVLLQVW--HHDGEC 244
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 241 LCPGPMGRITCEKACELHTFGRICKRCSQDQDQKSYVFLCPYQCCSCATGKMGLOQNE 300
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 245 VCPGFTGTCTGECACNFRGQSCQCCQCPGISGTRGLFCLLPYQCCSGSGSWRGSQOUE 304
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 301 ACHEGFGVSPKLRQSN--NN--EM--TQPPQVTLSPQWQSLQTEPESLPPKMLPKIVLPDHL 360
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 305 ACAPGPHGALVTRIGCCQDQDQDQDQDQDQDQDQDQDQDQDQDQDQDQDQDQDQDQD 361
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 361 VNSGKNPI--CKASGWPLPTNEEMTLVKPDGTVLHPKDFNHTDHFVSVAITFTHRLPPDS 419
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 362 FNLETMPRIN--AAAGNPPFPVPGSIELPKPGDGVLLSTKALIVEPEKTTAEFEVPLVLADS 421
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 420 GVWVCSVNTVAGWVKEKFNISVVLKPKLNAINVI 454
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 422 GPWECRVSTSGQSQSRKFKVNVKVPVPIAAPRL 456
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 12
US-08-469-537A-97
: Sequence 97, Application US/08469537A
: Patent No. 5843749
: GENERAL INFORMATION:
: APPLICANT: Maisonnier, et al.
: TITLE OF INVENTION: EHK AND POP TYPING
: NUMBER OF SEQUENCES: 107
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Regeneron Pharmaceuticals, Inc.
: STREET: 777 Old Saw Mill River Road
: CITY: Tarrytown
: STATE: NY
: COUNTRY: U.S.A.
: ZIP: 10591
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette

```



```

QY 121 QASFLPATLTMTVDKGNVNISFKKVLKEEDAVIYKNGSFTHSVPRRIEVPDILEVHLPH 180
DB 121 QASFLPATLTMTVDKGNVNISFKKVLKEEDAVIYKNGSFTHSVPRRIEVPDILEVHLPH 180

QY 181 AQPQAGWYSARYTGNLTFTSAFTLIVPRFAKWKWPEVNHILTA-MNNWV-HEITGEV 240
DB 181 AQPQAGWYSARYTGNLTFTSAFTLIVPRFAKWKWPEVNHILTA-MNNWV-HEITGEV 240

QY 241 LCPFGEMHPTTEKAKVPEHTEGRTCKEKESEFEKQKSEVFLPDYFGSCATWKKLQCN 300
DB 241 LCPFGEMHPTTEKAKVPEHTEGRTCKEKESEFEKQKSEVFLPDYFGSCATWKKLQCN 300

QY 301 ACHPGFYGDCKLKRCSCNNGHMCURFQGCUCSPQWGLQCEKEDIPRMIPKIVDLPHLE 360
DB 301 ACHPGFYGDCKLKRCSCNNGHMCURFQGCUCSPQWGLQCEKEDIPRMIPKIVDLPHLE 360

QY 361 VNSGKFNPIKASGWPPTNEEMTLVKPDGTVLKPDKFNHTDHSVAITTHILPDPDS 420
DB 361 VNSGKFNPIKASGWPPTNEEMTLVKPDGTVLKPDKFNHTDHSVAITTHILPDPDS 420

QY 421 VVWCVSVNTVAGWVEKPNISVKVLPKELNAPVDTGHNFAVINISSEYFG 472
DB 421 VVWCVSVNTVAGWVEKPNISVKVLPKELNAPVDTGHNFAVINISSEYFG 472

RESULT 2
protein-tyrosine kinase (EC 2.7.1.12), receptor type tek precursor bovine
N:Alternate names: receptor tyrosine kinase tie-2
C:Species: Bos primigenius tauris (cattle)
C>Date: 28-Oct-1995 #sequence_revision 01-Nov-1995 #text_change 16-Jul-1999
C:Accession: S57846; S32691
R:Sato, T.N.; Qin, Y.; Kozak, C.A.; Audus, K.L.
Proc. Natl. Acad. Sci. U.S.A. 90, 9355-9358, 1993
A:Title: tie-1 and tie-2 define another class of putative receptor tyrosine kinase genes
A:Reference number: S57845; MUII-04022374
A:Accession: S57846
A>Status: preliminary; nucleic acid sequence not shown, translation not shown
A:Molecule type: mRNA
A:Residues: 1-1125 <SA2>
A:Cross-references: EMBL: X71424; NID: q296577; PDB: 1AA50555.1; PDB: q296578
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1994
R:Sato, T.N.; Qin, Y.; Kozak, C.A.; Audus, K.L.
submitted to the EMBL Data Library, March 1993
A:Reference number: S32690
A:Accession: S32690
A:Molecule type: mRNA
A:Residues: 1-1125 <SA2>
A:Cross-references: EMBL: X71424; NID: q296577; PDB: 1AA50555.1; PDB: q296578
C:Superfamily: protein tyrosine kinase, receptor type tie; EGF homology; fibronectin
C:Keywords: ATP, autophosphorylation, diphosphatase, glycoprotein, phosphoprotein, phospho
E:1-22/Domain: signal sequence #status predicted <SIG>
E:27-104/Domain: protein tyrosine kinase, receptor type tek #status predicted <MAT>
E:145-137/Region: cell attachment (R-G-D) motif
E:211-251/Domain: EGF homology <EG1>
E:255-298/Domain: EGF homology <EG2>
E:302-340/Domain: EGF homology <EG3>
E:342-426/Domain: immunoglobulin homology <IM2>
E:447-527/Domain: fibronectin type III repeat homology <FN4>
E:542-626/Domain: fibronectin type III repeat homology <FN3B>
E:639-721/Domain: fibronectin type III repeat homology <FN3C>
E:753-774/Domain: transmembrane #status predicted <TM>
E:823-1100/Domain: protein kinase homology <KIN>
E:H31-R40/Region: protein kinase ATP-binding motif
E:140-158,399,438,464,560,597,650,692/Binding site: carbohydrate (Asn) (covalent) #statu
E:856,873,965/Active site: Lys, Glu, Asp #status predicted

Query Match 93.58; Score 2435; DB 1; Length 1125;
Best Local Similarity 93.28; Pred. No. 5e-161;
Matches 440; Conservative 12; Mismatches 20; Indels 0; Gaps 0;

```

```

QY 1 MDSLASLVLCVSVLLSGTVEGAMDLILNSLPLVSDAEISLTICIASGWHPHEPTIGR 60
DB 1 MDSLASLVLCVSVLLSGTVEGAMDLILNSLPLVSDAEISLTICIASGWHPHEPTIGR 60

QY 61 FEALMNHQDPLFTVQTQVTPPEWAKKVVWVKEKASKINCAVYFCEGVPFGAIRITMKMPQ 120
DB 61 FEALMNHQDPLFTVQTQVTPPEWAKKVVWVKEKASKINCAVYFCEGVPFGAIRITMKMPQ 120

QY 121 QASFLPATLTMTVDKGNVNISFKKVLKEEDAVIYKNGSFTHSVPRRIEVPDILEVHLPH 180
DB 121 QASFLPATLTMTVDKGNVNISFKKVLKEEDAVIYKNGSFTHSVPRRIEVPDILEVHLPH 180

QY 181 AQPQAGWYSARYTGNLTFTSAFTLIVPRFAKWKWPEVNHILTA-MNNWV-HEITGEV 240
DB 181 AQPQAGWYSARYTGNLTFTSAFTLIVPRFAKWKWPEVNHILTA-MNNWV-HEITGEV 240

QY 241 LCPFGEMHPTTEKAKVPEHTEGRTCKEKESEFEKQKSEVFLPDYFGSCATWKKLQCN 300
DB 241 LCPFGEMHPTTEKAKVPEHTEGRTCKEKESEFEKQKSEVFLPDYFGSCATWKKLQCN 300

QY 301 ACHPGFYGDCKLKRCSCNNGHMCURFQGCUCSPQWGLQCEKEDIPRMIPKIVDLPHLE 360
DB 301 ACHPGFYGDCKLKRCSCNNGHMCURFQGCUCSPQWGLQCEKEDIPRMIPKIVDLPHLE 360

QY 361 VNSGKFNPIKASGWPPTNEEMTLVKPDGTVLKPDKFNHTDHSVAITTHILPDPDS 420
DB 361 VNSGKFNPIKASGWPPTNEEMTLVKPDGTVLKPDKFNHTDHSVAITTHILPDPDS 420

QY 421 VVWCVSVNTVAGWVEKPNISVKVLPKELNAPVDTGHNFAVINISSEYFG 472
DB 421 VVWCVSVNTVAGWVEKPNISVKVLPKELNAPVDTGHNFAVINISSEYFG 472

RESULT 3
protein tyrosine kinase (EC 2.7.1.12), receptor type tek precursor mouse
N:Alternate names: protein-tyrosine kinase receptor type tie2
C:Species: Mus sp. (mouse)
C>Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 21-Jun-2000
C:Accession: 154237
R:Runtz, A.S.; Stackel, S.A.; Wilks, A.F.
Growth Factors 9, 99-105, 1993
A:Title: tie2, a putative protein tyrosine kinase from a new class of cell surface re
A:Reference number: 154237; MUII-94031116
A:Accession: 154237
A>Status: preliminary; translated from GH/EMBL/DDHJ
A:Molecule type: mRNA
A:Residues: 1-1125 <RES>
A:Cross-references: GB: S67051; NID: q452873; PDB: 1AB26683.1; PDB: q452874
C:Genetics:
A:Gene: tie2
C:Superfamily: protein tyrosine kinase, receptor type tie; EGF homology; fibronectin
C:Keywords: ATP, autophosphorylation, diphosphatase, glycoprotein, phosphoprotein, phospho
E:27-104/Domain: immunoglobulin homology <IMM>
E:211-251/Domain: EGF homology <EG1>
E:255-298/Domain: EGF homology <EG2>
E:302-340/Domain: EGF homology <EG3>
E:340-624/Domain: fibronectin type III repeat homology <3FR>
E:820-1097/Domain: protein kinase homology <FIN>
E:828-836/Region: protein kinase ATP binding motif

Query Match 92.48; Score 2407; DB 2; Length 1122;
Best Local Similarity 91.98; Pred. No. 4.3e-159;
Matches 434; Conservative 17; Mismatches 21; Indels 0; Gaps 0;

```


Db 61 FEALMNGHLEPLEVTCVTRFEWAKVVKWPEKASKINLAYPSPVPRQALPIPTMKMPQ 180
 QY 121 QASFLPATLITMTVDKGNVNISSKKVLKEELAVLYKNGSPHSHSVPRHEVPIILVHLPH 180
 Db 121 QASFLPATLITMTVDKGNVNISSKKVLKEELAVLYKNGSPHSHSVPRHEVPIILVHLPH 180
 QY 181 AQPDACVYSARYIGNLPTSAFTRIIVPRTEAKWKHRRNHILATAWNGWVHECTDGL 240
 Db 181 AQPDACVYSARYIGNLPTSAFTRIIVPRTEAKWKHRRNHILATAWNGWVHECTDGL 240
 QY 241 ICNCHMERTQKACDIEHTEKCHKPKROGCKSYVFLPRVQSSAPQWKLQDNE 300
 Db 241 ICNCHMERTQKACDIEHTEKCHKPKROGCKSYVFLPRVQSSAPQWKLQDNE 300
 QY 301 ACHPGYGHACILKRCNNMPEMTHKQGLKCSQWQGLQCEPEGRPMIPKIVLPDHL 360
 Db 301 ACHPGYGHACILKRCNNMPEMTHKQGLKCSQWQGLQCEPEGRPMIPKIVLPDHL 360
 QY 361 VNSGKFNPIKASGWPLPTNEMLVKKPCTVLHPTGNHTHFSVALPTIIPILPPSG 420
 Db 361 VNSGKFNPIKASGWPLPTNEMLVKKPCTVLHPTGNHTHFSVALPTIIPILPPSG 420
 QY 421 VWCVSNTVAGWKEKFNISVVKLPKPLNAPNVIDTGNFVAVINISSEPYG 472
 Db 421 VWCVSNTVAGWKEKFNISVVKLPKPLNAPNVIDTGNFVAVINISSEPYG 472

RESULT 4
 JN0712
 protein-tyrosine kinase (EC 2.7.1.112), receptor type tek precursor mouse
 N:Alternate names: protein-tyrosine kinase, receptor type tie-2
 C:Species: Mus musculus (house mouse)
 C:Date: 27 Jun-1994 #sequence_revision 27 Jun-1994 #text_change 16-Jul-1999
 C:Accession: JN0712; S43495; S43494; S33142
 R:Iwama, A.; Hanaguchi, I.; Hashiyama, M.; Murayama, Y.; Yasunaga, K.; Suda, T.
 Biochem. Biophys. Res. Commun. 195, 301-309, 1993
 A:Title: Molecular cloning and characterization of mouse TIE and TEK receptor tyrosine kinase genes
 A:Reference number: S57848; MIM:94022374
 A:Accession: S57848
 A:Status: preliminary; nucleic acid sequence not shown, translation not shown
 A:Molecule type: mRNA
 A:Residues: 1-786,788 1123-1123, SAT
 A:Cross-references: EMBL X71426, NID 9220439, PID 9220433
 A:Note: the nucleotide sequence was submitted to the EMBL data library, March 1993
 R:Dumont, D.J.
 submitted to the EMBL Data library, April 1993
 A:Reference number: S43495
 A:Accession: S43495
 A:Molecule type: mRNA
 A:Residues: 1-537,537,737-786,788-1123 1123-1123, QWQ
 A:Cross-references: EMBL X62553, NID 9220439, PID 9220433
 R:Dumont, D.J.; Yanaguchi, I.P.; Conlon, P.A.; Bossant, J.; Breitman, M.L.
 Oncogene 7, 1471-1480, 1992
 A:Title: tek, a novel tyrosine kinase gene located on mouse chromosome 4, is expressed in hematopoietic cells
 A:Reference number: S43494; MIM:92334855
 A:Accession: S43494
 A:Molecule type: mRNA
 A:Residues: 823-1123 <QWQ>
 A:Cross-references: EMBL X67553
 C:Comment: Receptor tyrosine kinase ligand systems play an important role in the constitutive activation of tyrosine kinases
 C:Genetics:
 A:Map position: 4
 C:Superfamily: protein-tyrosine kinase, receptor type tie, EGF homology: fibronectin type III, ATP, autophosphorylation, duplication, glycoprotein, phosphoprotein, phospho
 E:1-22/Domain: signal sequence status predicted <S>
 E:23-1123/Product: protein-tyrosine kinase, receptor type tie-2 status predicted <M>

F:12-104/Domain: immunoglobulin homology <IM>
 F:135-137/Region: cell attachment (R-G-D) motif
 F:211-251/Domain: EGF homology <EG>
 F:255-298/Domain: EGF homology <EG>
 F:302-342/Domain: EGF homology <EG>
 F:364-426/Domain: immunoglobulin homology <IM>
 F:447-455/Domain: fibronectin type III repeat homology <FN3>
 F:540-624/Domain: fibronectin type III repeat homology <FN3>
 F:637-719/Domain: fibronectin type III repeat homology <FN3>
 F:751-771/Domain: transmembrane status predicted <TM>
 F:821-1009/Domain: protein kinase homology <PK>
 F:829-837/Region: protein kinase ATP-binding motif
 F:140-158,499,444,528,545,546/Binding site: carbohydrate (Asn) (covalent) #st
 F:854,871,963/Active site: Lys, Glu, Asp #status predicted

Query Match 92.48; Score 2407; DR 1; Length 1123;
 Best Local Similarity 91.98; Pred. No. 4.4e-159;
 Matches 434; Conservative 19; Mismatches 21; Indels 0; Gaps 0;

QY 1 MSLASIVLCVSIIISIVECAMLEIIINSIPVSHAEISLICIASOWRHPHPIICRD 60
 Db 1 MSLASIVLCVSIIISIVECAMLEIIINSIPVSHAEISLICIASOWRHPHPIICRD 60
 QY 61 FEALMNGHLEPLEVTCVTRFEWAKVVKWPEKASKINLAYPSPVPRQALPIPTMKMPQ 120
 Db 61 FEALMNGHLEPLEVTCVTRFEWAKVVKWPEKASKINLAYPSPVPRQALPIPTMKMPQ 120
 QY 121 QASFLPATLITMTVDKGNVNISSKKVLKEELAVLYKNGSPHSHSVPRHEVPIILVHLPH 180
 Db 121 QASFLPATLITMTVDKGNVNISSKKVLKEELAVLYKNGSPHSHSVPRHEVPIILVHLPH 180
 QY 181 AQPDACVYSARYIGNLPTSAFTRIIVPRTEAKWKHRRNHILATAWNGWVHECTDGL 240
 Db 181 AQPDACVYSARYIGNLPTSAFTRIIVPRTEAKWKHRRNHILATAWNGWVHECTDGL 240
 QY 241 ICNCHMERTQKACDIEHTEKCHKPKROGCKSYVFLPRVQSSAPQWKLQDNE 300
 Db 241 ICNCHMERTQKACDIEHTEKCHKPKROGCKSYVFLPRVQSSAPQWKLQDNE 300
 QY 301 ACHPGYGHACILKRCNNMPEMTHKQGLKCSQWQGLQCEPEGRPMIPKIVLPDHL 360
 Db 301 ACHPGYGHACILKRCNNMPEMTHKQGLKCSQWQGLQCEPEGRPMIPKIVLPDHL 360
 QY 361 VNSGKFNPIKASGWPLPTNEMLVKKPCTVLHPTGNHTHFSVALPTIIPILPPSG 420
 Db 361 VNSGKFNPIKASGWPLPTNEMLVKKPCTVLHPTGNHTHFSVALPTIIPILPPSG 420
 QY 421 VWCVSNTVAGWKEKFNISVVKLPKPLNAPNVIDTGNFVAVINISSEPYG 472
 Db 421 VWCVSNTVAGWKEKFNISVVKLPKPLNAPNVIDTGNFVAVINISSEPYG 472

RESULT 5
 JN0771
 protein-tyrosine kinase (EC 2.7.1.112), receptor type tek precursor mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 10 Jun-1993 #sequence_revision 10 Jun-1993 #text_change 21-Jul-2000
 C:Accession: JN0771
 R:Horita, K.; Yagi, T.; Kohmura, N.; Tomooka, Y.; Ikawa, Y.; Aizawa, S.
 Biochem. Biophys. Res. Commun. 189, 1747-1753, 1992
 A:Title: A novel tyrosine kinase, tek, expressed in murine embryonic stem cells.
 A:Reference number: JN0771; MIM:93129253
 A:Accession: JN0771
 A:Molecule type: mRNA
 A:Residues: 1-1125 <RP>
 A:Cross-references: GR P13738, NID 9220439, PID 9220433, PID 9220439
 C:Genetics:
 A:Gene: tek
 C:Superfamily: protein-tyrosine kinase, receptor type tek, EGF homology: fibronectin
 C:Keywords: ATP, autophosphorylation, duplication, glycoprotein, phosphoprotein, phospho
 E:1-22/Domain: signal sequence status predicted <S>
 E:23-1125/Product: protein-tyrosine kinase, receptor type tek status predicted <M>

F:47-104/Domain: immunoglobulin homology <IM1>
 F:135-137/Region: cell attachment (P-G-D) motif
 F:212-252/Domain: EGF homology <EGF>
 F:256-269/Domain: EGF homology <EGF>
 F:303-341/Domain: EGF homology <EGF>
 F:365-427/Domain: immunoglobulin homology <IM2>
 F:448-526/Domain: fibronectin type III repeat homology <FN3A>
 F:541-625/Domain: fibronectin type III repeat homology <FN3B>
 F:638-720/Domain: fibronectin type III repeat homology <FN3C>
 F:752-773/Domain: transmembrane status predicted <TM>
 F:824-1100/Domain: protein kinase homology <KIN>
 F:831-849/Region: protein kinase ATP-binding motif
 F:140-158,400,439,465,559,649,691/Binding site: carbohydrate (Asn) (covalent) *status predicted
 F:856,873,965/Active site: Lys, Glu, Asp *status predicted

Query Match 89.7% Score 2337.5 DH 1 Length 1126
 Best Local Similarity 89.9% Pred. No. 2.9e-154
 Matches 426 Conserved 14 Mismatches 26 Indels 3 Gaps 2

QY 1 MDSIASVLCVSLISGTVHGMILLINSPLVSDAFISLITCIASGWPDPHPPTIGRD 60
 DQ 1 MDSIAGVLCVSLISGTVHGMILLINSPLVSDAFISLITCIASGWPDPHPPTIGRD 60
 QY 61 FEAALNQDDPLEVTQGVTFWAKKVVWKKPEKSKINAYTFSPVSRREAITPTMKMPG 120
 DQ 61 FEAALNQDDPLEVTQGVTFWAKKVVWKKPEKSKINAYTFSPVSRREAITPTMKMPG 120
 QY 121 QASFLPATITWVKGDNVNISSPKKVLIKEDAVIVKNGSFH--SVPRHVPDILEVH 178
 DQ 121 QASFLPATITWVKGDNVNISSPKKVLIKEDAVIVKNGS--LPLSAPGMKYLDILEVH 179
 QY 179 FPAQDQACVYSARYIGCNLTPTSAFTPLIVRCEAKWQPCNHLCTACMNGNVCHEDTG 238
 DQ 180 FPAQDQACVYSARYIGCNLTPTSAFTPLIVRCEAKWQPCNHLCTACMNGNVCHEDTG 239
 QY 239 EITTFETEMHPTKAKVLEHTFPTWEPSSGDEPKSVVDFDFYGSATGKGLG 298
 DQ 240 EITTFETEMHPTKAKVLEHTFPTWEPSSGDEPKSVVDFDFYGSATGKGLG 299
 QY 299 FPAQDQACVYSARYIGCNLTPTSAFTPLIVRCEAKWQPCNHLCTACMNGNVCHEDTG 358
 DQ 300 NEACPSSGYYGDPCKLRGCTINEEDREFGQGLCSGQWGLDCEKEGKRWTPU 359
 QY 359 FVNSGKFNPLCKASGWPDPHPPTIGRDPVLPKDPNHDHVSVAIFTHRIILPD 418
 DQ 360 FVNSGKFNPLCKASGWPDPHPPTIGRDPVLPKDPNHDHVSVAIFTHRIILPD 419
 QY 419 SGVWVCSVNTVAGVMEKPEKPNISVKVLKPLNAPNVI 472
 DQ 420 SGVWVCSVNTVAGVMEKPEKPNISVKVLKPLNAPNVI 473

RESULT 6
 S57845
 protein-tyrosine kinase (E0 2 7 1 112), receptor type tie precursor - bovine
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 28-Oct-1995 #sequence-revision 03-Nov-1995 #text change 16-Jul-1999
 C:Accession: S57845; S32640
 R:Sato, T.N.; Glin, Y.; Kozak, C.A.; Audus, K.L.
 Proc. Natl. Acad. Sci. U.S.A. 90 9455-9458 1993
 A:Title: tie-1 and tie-2 define another class of putative receptor tyrosine kinase genes
 A:Reference number: S57845; MUID:94022374
 A:Accession: S57845
 A:Status: preliminary, nucleic acid sequence not shown, translation not shown
 A:Molecule type: mRNA
 A:Residues: 1-1116 <SAT>
 A:Cross-references: EMRL: X71474; NID: 9296575; PIRN: AA0554 1; PFD: 429457
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, March 1993
 C:Superfamily: protein tyrosine kinase, receptor type tie, EGF homology, fibronectin type
 C:Keywords: ATP, autophosphorylation, duplication, dyad-symmetry, phosphoprotein, phospho
 F:1-21/Domain: signal sequence *status predicted <SIG>
 F:22-1116/Protein: protein tyrosine kinase, receptor type tie *status predicted <MAT>

F:36-108/Domain: immunoglobulin homology <IM1>
 F:213-253/Domain: EGF homology <EGF>
 F:257-303/Domain: EGF homology <EGF>
 F:304-342/Domain: EGF homology <EGF>
 F:363-426/Domain: immunoglobulin homology <IM2>
 F:447-528/Domain: fibronectin type III repeat homology <FN3A>
 F:540-631/Domain: fibronectin type III repeat homology <FN3B>
 F:646-728/Domain: fibronectin type III repeat homology <FN3C>
 F:753-784/Domain: transmembrane status predicted <TM>
 F:835-1112/Domain: protein kinase homology <KIN>
 F:843-851/Region: protein kinase ATP-binding motif
 F:43-106,376-424/Disulfide bonds: *status predicted
 F:84,150,501,594,707/Binding site: carbohydrate (Asn) (covalent) *status predicted
 F:868,885,977/Active site: Lys, Glu, Asp *status predicted

Query Match 33.5% Score 873.5 DH 1 Length 1146
 Best Local Similarity 39.3% Pred. No. 9.4e-53
 Matches 176 Conserved 16 Mismatches 176 Indels 6 Gaps 12

QY 16 TSGTVHGMILLINSPLVSDAFISLITCI-----ASGWRPHEPTIGRDFEALNN 66
 DQ 15 LASNVGAADVILLALVLELPDPPEFLTCVSCACACRGSDAWCP--PILLEKDPHIVPT 72
 QY 67 QH--QPLEVTQGVTFWAKKVVWKKPEKSKINAYTFSPVSRREAITPTMKMPG 123
 DQ 73 PRWQPP-----HIAKNGSSRIVRGFSQPSDILGVFSCVG---GGTRVLYVHNSGAAH 124
 QY 124 FLPATITWVKGDNVNISSPKKVLIKEDAVIVKNGSFHSHVPRHVPD---LLEVHILPD 180
 DQ 125 LLDKVTHTVNGDQAVLSARVKEKQTDVWKSNGSYFTIDIEHQDQDQLQ--LPR 182
 QY 181 APODQACVYSARYIGCNLTPTSAFTPLIVRCEAKWQPCNHLCTACMNGNVCHEDTG 240
 DQ 183 VDPSSGTSYATVPEASHISAPFPLIVRCEAKWQPCNHLCTACMNGNVCHEDTG 242
 QY 241 LPPHFMPT--EKAKVLEHTFPTWEPSSGDEPKSVVDFDFYGSATGKGLG 300
 DQ 243 VPPHFGTIFCEQACPKSRFGVSCVQCPGISGCRGLIFLPPYGCSCGSCGSCGSC 302
 QY 301 ACHPEKAPKPKKESNNGRPMRFGQGLCSGQWGLDCEKEGKRWTPU 360
 DQ 303 A-APKPFAC-HLQVQUNGT--DPSSQVWPSAHGMHCEKSD---RIPQILDWYSELE 359
 QY 361 VNSGKFNPLCKASGWPDPHPPTIGRDPVLPKDPNHDHVSVAIFTHRIILPD 418
 DQ 460 FNLDITMPKINCAAGNPPVWGSMEILKKPKDITVLSKATVEPDR--TTAFEPVPLAID 418
 QY 419 SGVWVCSVNTVAGVMEKPEKPNISVKVLKPLNAPNVI 454
 DQ 419 SGLWPCVSVSGGQSRPRINVKVDPVPLAID 454

RESULT 7
 S24066
 protein-tyrosine kinase (E0 2 7 1 112), receptor type tie precursor - human
 C:Species: Homo sapiens (man)
 C:Date: 19-Feb-1994 #sequence-revision 11-Aug-1995 #text change 21-Jul-2000
 C:Accession: S24066; C38269; 152613
 R:Partanen, J.; Armstrong, E.; Maekelae, T.P.; Korhonen, J.; Sandberg, M.; Penkonen,
 Mol. Cell Biol. 12 1668-1707 1992
 A:Title: A novel endothelial cell surface receptor tyrosine kinase with extracellular
 A:Reference number: S24066; MUID:92195316
 A:Accession: S24066
 A:Molecule type: mRNA
 A:Residues: 1-1138 <PAR>
 A:Cross-references: EMBL: X60957
 P:Partanen, J.; Maekelae, T.P.; Allitalo, E.; Leheslahti, H.; Allitalo, K.
 Proc. Natl. Acad. Sci. U.S.A. 87 8913-8917 1990
 A:Title: Putative tyrosine kinases expressed in K-562 human leukemia cells.
 A:Reference number: A38268; MUID:91062389
 A:Accession: C38269
 A:Status: preliminary, nucleic acid sequence not shown, not compared with conceptual

A:Molecule type: mRNA
 A:Residues: 981-1034 <RAW>
 A:Experimental source: clone JTK14
 R:Korhonen, J.; Lahtinen, I.; Halmekyto, M.; Alhonen, L.; Jänne, J.; Dumont, D.; Allitala
 Blood 86, 1828-1835, 1995
 A:Title: Endothelial specific gene expression directed by the tie gene promoter in vivo.
 A:Reference number: 152613, MUID:95383653
 A:Accession: JN0711, S33141; S57847; A48926; 165403
 R:Iwama, A.; Hamaguchi, I.; Hashiyama, M.; Mura, Y.; Yasunaga, K.; Suda, T.
 Biochem. Biophys. Res. Commun. 195, 301-309, 1993
 A:Title: Molecular cloning and characterization of mouse TIE and TEK receptor tyrosin
 A:Reference number: JN0711, MUID:93371421
 A:Accession: JN0711
 C:Genetics:
 A:Gene: GIM:TIE; JTK14
 A:Cross-references: GIM:212973, GIM:600222
 A:Map position: lp34-1p33
 C:Function:
 A:Description: Catalyzes the phosphorylation of a peptidyl tyrosine residue by ATP
 C:Superfamily: Protein-tyrosine kinase, receptor type tie, EGF homology, fibronectin tyb
 C:Keywords: ATP, autophosphorylation, duplication, glycoprotein, phosphoprotein, phospho
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-1138/Domain: protein tyrosine kinase, receptor type tie #status predicted <MAT>
 F:36-107/Domain: immunoglobulin homology <IM1>
 F:215-255/Domain: EGF homology <EG1>
 F:259-302/Domain: EGF homology <EG2>
 F:306-344/Domain: EGF homology <EG3>
 F:365-428/Domain: immunoglobulin homology <IM2>
 F:449-530/Domain: fibronectin type III repeat homology <FN1A>
 F:542-633/Domain: fibronectin type III repeat homology <FN1B>
 F:642-730/Domain: fibronectin type III repeat homology <FN1C>
 F:761-786/Domain: transmembrane #status predicted <TM>
 F:837-1114/Domain: protein kinase homology <KIN>
 F:845-853/Region: protein kinase ATP-binding motif
 F:43-105,372-426/Disulfide bonds: #status predicted
 F:83,161,503,596,709/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:870,887,979/Active site: Lys, Glu, Asp #status predicted

Query Match 33.2%, Score 865.5, DB 1, Length 1138;
 Best Local Similarity 38.2%, Pred. No. 3.4e 52;
 Matches 174, Conservative 71, Mismatches 185, Indels 25, Gaps 4;
 QY 14 LLLSGTVEGAMDLILNLSPLVSDATSLTCTI-----ASGQFHEFTITIGDFFAL 64
 DB 13 LFLASHVGAAGVETLLANRLTLPQFFLTGVSAGASGASDAP--P--PLLEKLDIV 70
 QY 65 MNHQDPLEVTQVTRWAKKVKWKR--EKASKINAYFPGKVPGBAIPRTMKMPQAS 123
 DB 71 FTFPGTFLF----IAPNCSHQVLESEFKFSGLVGVFSCVQAGAKRIPVIVVHNSGCAH 126
 QY 124 FLPATLTMTVDKGNVNISEKKVLIKEDNAVYKNGSFTHSVPRIEVPD---LLEVLPH 180
 DB 127 LLPDKVTHVTKGDTAVLSARVUIKOTDVIKNSGSIPTTLDWIEAQQGRELQ--LPN 184
 QY 181 AQPDAQVYSARYIGGNIPTSAFTPIIVPQCPACKKCPQNHICTACMNGVCHDITGPC 240
 DB 185 VQPPSSGIYSATVLEASPLGSAFFELIVRGAGPWPQVPTKEDPGLHGVWCHIDPQV 244
 QY 241 ICPFGWGRTRTEKATELHTFGTPTKEP--SQEGLKSKYVFLEDPYPSATLWKLQYNE 400
 DB 245 VTPQPGTPTHTFGACAPGCPQSGQWQVQVLSQCPGLPGLDNYGDSGSGWQSGQCP 304
 QY 301 AHPFGVGPWKLKPS--NNGCEM--TQFPGVPLSHWQWILQ--TEREIPPMPTFKVLPQHE 400
 DB 305 AAFQHPGAPRLQ--LQ--NNGVLTQPSG--VQESWIDVLTKEK----PFPQLNMASELE 401
 QY 361 VNSGKNPI--CKASGWLPIPIHEMTLVKPDGTVLHPKDFNHTDHSVAIFTHRIILPPDS 419
 DB 362 FLETPRINCAAGNFPVVRGSEIELRKPDGTVLLSTKAIVEPEKTTAEFEVPRVLVADS 421
 QY 420 GVWCVSVNTVAGVVEKPFNLISVKVLPKPLNAPVI 454
 DB 422 GPWECRVSTGGQDSRRKPKVNVKVPVPLAAPRL 456

RESULT 8

JN0711
 C:Species: Mus musculus (house mouse)
 C:Date: 14 Jul 1994 #sequence_revision 14 Jul 1994 #text_change 16-Jul-1999
 C:Accession: JN0711; S33141; S57847; A48926; 165403
 R:Iwama, A.; Hamaguchi, I.; Hashiyama, M.; Mura, Y.; Yasunaga, K.; Suda, T.
 Biochem. Biophys. Res. Commun. 195, 301-309, 1993
 A:Title: Molecular cloning and characterization of mouse TIE and TEK receptor tyrosin
 A:Reference number: JN0711, MUID:93371421
 A:Accession: JN0711
 A:Molecule type: mRNA
 A:Residues: 1-1134 <RAW>
 A:Cross-references: GR:X73060; NID:9402601; PDB:CAA52148.1; PDB:9402602
 R:Sato, T.N.; Oin, Y.; Kozak, C.A.; Andus, K.L.
 submitted to the EMBL data library, March 1993
 A:Accession: S33141
 A:Molecule type: mRNA
 A:Residues: 1-598, 1-1500-1134 <SAT>
 A:Cross-references: EMBL:X731425; NID:9495610; PDB:CAA50556.1; PDB:9495611
 R:Sato, T.N.; Oin, Y.; Kozak, C.A.; Andus, K.L.
 Proc. Natl. Acad. Sci. U.S.A. 90, 9355-9358, 1993
 A:Title: tie 1 and tie 2 define another class of putative receptor tyrosine kinase ge
 A:Reference number: S57845, MUID:94023374
 A:Accession: S57847
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: mRNA
 A:Residues: 1-598, 1-1500-1134 <SAT>
 A:Cross-references: EMBL:X71425, NID:9495610, PDB:CAA50556.1, PDB:9495611
 R:Korhonen, J.; Partanen, J.; Armstrong, E.; Vahdokari, A.; Elenius, K.; Jalkanen, M
 Blood 80, 2548-2555, 1992
 A:Title: Enhanced expression of the tie receptor tyrosine kinase in endothelial cells
 A:Reference number: A48926, MUID:93043301
 A:Accession: A48926
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 221-527,740-890 <RG>
 A:Note: sequence extracted from NCHI backbone (NCBI:P118660, NCBI:P118662)
 R:Korhonen, J.; Lahtinen, I.; Halmekyto, M.; Alhonen, L.; Jänne, J.; Dumont, D.; Alit
 Blood 86, 1828-1835, 1995
 A:Title: Endothelial-specific gene expression directed by the tie gene promoter in vi
 A:Reference number: 152613, MUID:95383653
 A:Accession: 155403
 A:Status: preliminary; translated from GR/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-19 <RES>
 A:Cross-references: GR:S79346; NID:91086020
 C:Superfamily: protein-tyrosine kinase, receptor type tie; EGF homology; fibronectin
 C:Keywords: ATP, autophosphorylation, duplication, glycoprotein, phosphoprotein; phos
 F:1-21/Domain: signal sequence #status predicted <SIG>
 F:22-1134/Domain: protein tyrosine kinase, receptor type tie #status predicted <MAT>
 F:36-105/Domain: immunoglobulin homology <IM1>
 F:214-254/Domain: EGF homology <EG1>
 F:257-300/Domain: EGF homology <EG2>
 F:304-342/Domain: EGF homology <EG3>
 F:363-426/Domain: immunoglobulin homology <IM2>
 F:447-528/Domain: fibronectin type III repeat homology <FN1A>
 F:540-629/Domain: fibronectin type III repeat homology <FN1B>
 F:648-735/Domain: fibronectin type III repeat homology <FN1C>
 F:757-782/Domain: transmembrane #status predicted <TM>
 F:43-105/Domain: protein kinase homology <KIN>
 F:841-849/Region: protein kinase ATP-binding motif
 F:43-103,370-424/Disulfide bonds: #status predicted
 F:81,159,501,592,705/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:866,883,975/Active site: Lys, Glu, Asp #status predicted

Query Match 31.5%, Score 820.5, DB 1, Length 1134;
 Best Local Similarity 36.0%, Pred. No. 4.5e-49;

Matches 169; Conservative 70; Mismatches 191; Indels 39; Gaps 7;

QY 5 ASLVAGVSLLSGIVEGAMLLILNLSPLVSDAETSILGIA-----SGWRPHEPITIGRD 60
 DB 6 SSLL--PTLEASIVAGSVOLITLANLITDPORFLLTGVSGEASAGRSSDPLLEKD 63
 QY 61 -----PIALMNQHDPLEVTDVTRHWAKKVVVKRKKASKINCAVEFGGRVR 107
 DB 64 DRIVEPFGPGPYLARNSSHQVTLF-----GFSKPSDLVGVSCVGGAG 108
 QY 108 GEAIRTRMKRQASFLPATLITVQKIGNVINISFKKVLKEEDAVIYKNGSFIHSVR 167
 DB 109 AKRTKLVVHNSIPGAHLPPDKVTHVKNQGLVIAVSAHVHAKGLDIVLWKNNGSYFNIJW 168
 QY 168 HEVPE--IIEVHPIACFOVAVY--AFYVGNELTSAFTRIIVRPGAKKWFTECNILCTA 226
 DB 169 QEADDPFQLGLQNLVPESSGLYATVIEASAFERLIVRQVAGRWLPQVVKIETG 228
 QY 227 QNNVCVCHEDICDCTEPCWCRICRATGHTGSGKPKWQVQKNSVYVFTIDVQSG 286
 DB 229 CLHGAWHIFHSEVWHPSEHGFHGFAGWPRRPGSGQWTPGIAHAPGLTFLFPQYS 288
 QY 287 FSTATGKGLDLENAHSEVSEGLKLESGNNEMTUPQGLDLSGQWASLWTEEGIP 346
 DB 289 CSCGSCWRCSCQACAPDHPGALHRCQCCQNGCICDPSGCVSGWGVGCEKSD-- 346
 QY 347 RMTPKIVDLPDIIHVNKGENEIKCKAGWPLPTNEEDTLVKPDGTVLHPKDFNHTDIFS 405
 DB 347 -RIPQILSMATEVEENIGTMRPNIAAGINPFPVRSKLRKPDGIMLLSTKVIPEPDT 405
 QY 406 VALPTIHKLLPDSGVVCSVNTVAGWVEKFNISVVLKPKPLNAVNI 454
 DB 406 TAEFEVPSLILGDSGFWRVSTSRQSRPRFKVNVKVPVPLTAPPLL 454

RESULT 9
 T27283
 hypothetical protein Y64G10A.f - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T27283
 R:Ainscough, R.
 submitted to the EMBL Data Library, September 1999
 A:Reference number: 220336
 A:Accession: T27283
 A>Status: preliminary; translated from GR/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-1620 <WILL>
 A:Cross-references: EMBL:AI10498; NID:G1542403; PIDN:CAH54471.1; CESP:Y64G10A.f
 A:Experimental source: clone Y64G10A
 C:Genetics:
 A:Gene: CESP:Y64G10A.1
 A:Introns: 77/1: 116/1: 148/1: 282/1: 305/1: 425/1: 466/1: 548/1: 604/1: 661/2: 695/1: 7

Query Match 12.38; Score 321; DB 2; Length 1620;
 Best Local Similarity 34.48; Pred. No. 2.8e-14;
 Matches 69; Conservative 22; Mismatches 66; Indels 32; Gaps 7;

QY 211 CCAQAWD--TRND--TAMNNGVWTHGSGPGLDQCHMCPKATTHHPPCKPKWNSG 270
 DB 957 CPTGREGI KMGLG--KQUNCAICUINSQSGMCAWGSKKAKACAGLFGKQSKKCD- 1014
 QY 271 QPGKSYVETLEFYSFATQWGLA--NEA--HPGFYFPA--KLP--S--NNE--EM--TUPG-- 429
 DB 1015 ---CAQMHCHDSGRCGLHPPGKAGKRCDEKPSGLFGACCKGKSCQNGATCDSTVSGC 1071
 QY 330 LPSGWOGLQCRERGDPMTPKIVDLPDIEVNSGKPK-----NPICKASGPLPTNFMEL 385
 DB 1072 ECRPWGKKKCDR-----PCPD-----GREGECNAICDCT-----TINDTSM 1109
 QY 386 VKP 388
 DB 1

DB 1110 YNP 1112

RESULT 10
 T13954
 MEGF6 protein - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
 C:Accession: T13954
 R:Nakayama, M.; Nakajima, D.; Nagase, T.; Nomura, N.; Seki, N.; Ohara, O.
 Genomics SL, 27-34, 1998
 A:Title: Identification of high molecular-weight proteins with multiple EGF-like motif
 A:Reference number: Z14126; MUID:98360089
 A:Accession: T13954
 A>Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: mRNA
 A:Residues: 1-1574 <NAK>
 A:Cross-references: EMBL:AB011532; NID:G3449293; PIDN:BAA32462.1; PID:G4449294
 A:Experimental source: Strain Sprague-Dawley; brain
 C:Genetics:
 A:Gene: MEGF6

Query Match 12.28; Score 317.5; DB 2; Length 1574;
 Best Local Similarity 35.08; Pred. No. 4.8e-14;
 Matches 57; Conservative 21; Mismatches 62; Indels 24; Gaps 4;

QY 211 CEAKWKIPEDCNILCTAMNNGVWTHGSGPGLDQCHMCPKATTHHGRICKERCSG 270
 DB 777 CPGRWGLGCOFICACACGASGNPETHCICLQCFVGSRCQDTCSAGWYGTGQIRCA 845
 QY 271 QPGKSYVETLEFYSFATQWGLA--NEA--HPGFYFPA--KLP--S--NNE--EM--TUPG-- 428
 DB 836 ---CANDGH--DPTTGR--SAPAWTSL--SQEAVTSGHW--PECLHP--N--SA--HGN--DAVSGIL 862
 QY 329 CLCSQMOGLQCRERGDPMTPKIVDLPDIEVNSGKPKFNPICK 371
 DB 893 CLCAGYFGPHCRQ-----SCQGYGYGPGSE 918

RESULT 11
 T26972
 hypothetical protein Y47H9C.4 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000
 C:Accession: T26972
 R:Harris, B.
 submitted to the EMBL Data Library, October 1998
 A:Reference number: 220593
 A:Accession: T26972
 A>Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-1111 <WILL>
 A:Cross-references: EMBL:AI032657; PIDN:CAA21739.1; GSPDH:GN00019; CESP:Y47H9C.4
 A:Experimental source: clone Y47H9C
 C:Genetics:
 A:Gene: CESP:Y47H9C.4
 A:Map position: 1
 A:Introns: 50/2: 84/3: 158/3: 248/3: 437/4: 777/4: 811/2: 847/2: 1017/1: 1084/1
 C:Superfamily: classed as ankyrin repeat proteins; ankyrin repeat homology; EGF homol

Query Match 10.78; Score 280; DB 2; Length 1111;
 Best Local Similarity 27.39; Pred. No. 1.3e-11;
 Matches 73; Conservative 29; Mismatches 112; Indels 48; Gaps 10;

QY 132 TVDKGDVNNISFKKVLKEEDAVIYKNGSFIHSVPHEVHD-----IIFVHPIHAC 182
 DB 81 TVEKPGQ--KASYQQLVKKEKYVKQCGDGYQTKDHFCLPDCNPPCKKSKGLEPKKCED 149
 QY 183 PQDAGVVSAPYIGENLFTSAFTLIVRPGAKKWFTECNILCTAMNNGSYFNIJW 242
 DB 140 PCYGGKYCA-----SSCSVGTWGLQCSASCD--QINGANCDPEIAGLCP 181

QY 302 CHFGYGFDECKLRSON NEMDQDFQK QVSPQWQGLQCFER EGIPTMTK 351
 DB 552 CQGVGFQGLCKRF-CNCIDIMGTEGGWCQDTTCQCLCKKCFACQKDCQCDIAFYGYPPCKAC 610
 QY 352 IVLQDHLIFVNSCKFNPICKASOWPLPTNEEM 384
 DB 611 ACLE-----GAGTTSPECDAISQCEPCNGNFI 636
 RESULT 15
 S18253
 laminin alpha-1 chain precursor - fruit fly (Drosophila melanogaster)
 C:Species: Drosophila melanogaster
 C:Date: 16-Sep-1992 #sequence_revision 24-Jul-1997 #text_change 10-Dec-1999
 C:Accession: S28395; S18253
 R:Kusilek-Gallibet, M., Garrison, K., MacKrell, A.J., Fessler, J.H.
 BMOJ J. 11, 4519-4527, 1992
 A>Title: Laminin A chain: expression during Drosophila development and genomic sequence.
 A:Reference number: S28399; MUID:93049203
 A:Accession: S28395
 A:Status: preliminary
 A:Molecule type: nucleic acid
 A:Residues: 1-4712 <KUS>
 A:Cross-references: GR-906498; NID-q157799; PIDN:AAA-8663.1; PID:q157800
 R:Garrison, K., MacKrell, A.J., Fessler, J.H.
 J. Biol. Chem. 266, 22899-22904, 1991
 A>Title: Drosophila laminin A chain sequence, interspecies comparison, and domain structure
 A:Reference number: S18253; MUID:92078147
 A:Accession: S18253
 A:Molecule type: mRNA
 A:Residues: 1762-3712 <GAR>
 A:Cross-references: EMBL:M75882; NID:q157797; PIDN:AAA-28661.1; PID:q157798
 C:Genetics:
 A:Gene: Flybase: lana
 A:Cross-references: Flybase:FBtr0092526
 C:Superfamily: laminin alpha-1 chain; laminin G repeat homology; laminin-type EGF-like homology
 C:Keywords: basement membrane; cell binding; coiled coil; disulfide bond; extracellular
 F:273-330/Domain: laminin-type EGF-like homology <LE3>
 F:333-400/Domain: laminin-type EGF-like homology <LE3>
 F:541-584/Domain: laminin-type EGF-like homology <LE3>
 F:1776-2115/Domain: III <DOM3>
 F:1776-1806/Domain: laminin-type EGF-like homology #status atypical <LE1>
 F:1809-1856/Domain: laminin-type EGF-like homology <LE2>
 F:1859-1914/Domain: laminin-type EGF-like homology <LE3>
 F:1917-1967/Domain: laminin-type EGF-like homology <LE4>
 F:1970-2014/Domain: laminin-type EGF-like homology <LE5>
 F:2017-2061/Domain: laminin-type EGF-like homology <LE6>
 F:2044-2109/Domain: laminin-type EGF-like homology <LE6>
 F:2116-2697/Domain: I/II, Reptad repeats <DOM2>
 F:2698-3712/Domain: G <DOM5>
 F:2698-2863/Domain: repeat G1 <RG1>
 F:2864-3048/Domain: repeat G2 <RG2>
 F:3049-3223/Domain: repeat G3 <RG3>
 F:3079-3260/Domain: laminin G repeat homology <LG3>
 F:3334-3528/Domain: repeat G4 <RG4>
 F:3529-3712/Domain: repeat G5 <RG5>
 F:1847,1850,1943,2024,2196,2215,2267,2301,2323,2492,2524,2538,2569,2699,2720,2890,2938,3

Query Match 6.1k; Score 212; DB 2; Length 3712;
 Best Local Similarity 24.6k; Pred. No. 2.5e-06;
 Matches 70; Conservative 38; Mismatches 87; Indels 90; Gaps 18;

QY 210 KCEAKKMG PRCHNLCYACMNGVCHDICEGICDPCGPMGRTCKKACELHTFCRTCKKRC 268
 DB 1407 KCRSYVGFDPCKP--CKCPNSAMCEPTTGECPMPNVIGDLGER-CAPTYG-----F 1457
 QY 269 SQGECKSVFCLDPQYK-----GSCATGMKGLQCNFACHGFGV-PDCKL 313
 DB 1458 HOVLQCEE--CAQNPWCIANGNSQICHPNGICHCQNIHGRAC-IVCSNGVFNPHCE- 1512
 QY 314 KCSNNG---EMCDREFG-CLASPWUQLOCE-----REGIP----- 346

EP 1513 QCSCHKPELLEFV-LKILHGA-FYKKNVVGPQCGVGTNLLNESNPDGYTTFCPEKTS 1472
 QY 347 -----RMTPKPIVDLPDHLFVNSGKFN-PICKASOWPLPTNEEM---FLKPKDKIVLHPKD 397
 DB 1573 RCD SAYLHVYVNSLKKHVSITITPEFHESIKFDMWVVADELLNFILKADPTREVD 1632
 QY 398 -----FNUTDHFVSV---AIFTIH--RILPPD 418
 DB 1633 ERPAVFGVIDYLLNCGNNHISAYGGHAYTLHFTSGPDGKYIVAH 1677

Search completed: September 23, 2002, 23:15:22
 Job time: 138 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd

OM protein - protein search, using sw model

Run on: September 23, 2002, 22:15:24, Search time 17.75 seconds
(without alignments)
1029.613 Million cell updates/sec

Title: US-09-733-764-2_COPY_1_472

Perfect score: 2605

Sequence: 1 MUSLASLVICGVSLILISGV.....VIDTGHNFVNISSEPYFG 472

Scoring table:
Gapop 10.0, Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 20300000000

Post processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_40*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES					Description
Result No	Score	Match	Length	ID	
1	2505	100.0	1124	1 TIE2_HUMAN	Q62763 homo sapien
2	2435	93.5	1125	1 TIE2_BOVIN	Q06807 bos taurus
3	2407	92.4	1122	1 TIE2_MOUSE	Q02858 mus musculus
4	873.5	33.5	1136	1 TIE1_BOVIN	Q06805 bos taurus
5	865.5	33.2	1138	1 TIE1_HUMAN	P35590 homo sapien
6	820.5	31.5	1134	1 TIE1_MOUSE	Q06806 mus musculus
7	289	11.1	830	1 SREC_HUMAN	Q14162 homo sapien
8	231.5	8.9	685	1 DLL4_HUMAN	Q9NF61 homo sapien
9	221	8.5	3672	1 LM2_CAEEL	Q21313 caenorhabdi
10	212	8.1	3712	1 LMADROME	Q00174 drosophila
11	206.5	7.9	1964	1 NTC4_MOUSE	P31895 mus musculus
12	193.5	7.7	686	1 DLL4_MOUSE	Q61171 mus musculus
13	193.5	7.7	1408	1 SERR_DROME	P18168 drosophila
14	198.5	7.6	379	1 WIF1_HUMAN	Q9Y5W5 homo sapien
15	198.5	7.6	833	1 D1_DROME	P10041 drosophila
16	198	7.5	723	1 DLL1_HUMAN	Q03548 homo sapien
17	195.5	7.5	379	1 WIF1_MOUSE	Q06548 mus musculus
18	193	7.4	2444	1 NTC1_HUMAN	P46531 homo sapien
19	192.5	7.4	722	1 DLL1_MOUSE	Q61483 mus musculus
20	192.5	7.4	1429	1 I112_CAEEL	P14585 caenorhabdi
21	192	7.4	2635	1 LMA5_HUMAN	O15230 homo sapien
22	191.5	7.4	2524	1 NOTC_XENLA	P21783 xenopus lae
23	191.5	7.4	3718	1 LMA5_MOUSE	Q61001 mus musculus
24	190.5	7.3	714	1 DLL1_RAT	P97677 rattus norv
25	189	7.3	473	1 FP2_MYTCA	Q25464 mytilus gal
26	187.5	7.2	3318	1 NTC3_MOUSE	Q61982 mus musculus
27	185.5	7.1	2703	1 NOTC_DROME	P07207 drosophila
28	184.5	7.1	1808	1 TENA_CHICK	P10039 gallus gal
29	184	7.1	1295	1 GEP1_CAEEL	P13508 caenorhabdi
30	183	7.0	2481	1 UN52_CAEEL	Q06561 caenorhabdi
31	183	7.0	2531	1 NTC1_MOUSE	Q01705 mus musculus
32	182	7.0	2531	1 NTC1_RAT	Q07008 rattus norv
33	181.5	7.0	2437	1 NOTC_HHARE	P46530 brachydanio

RESULT 1

ID	TIE2_HUMAN	STANDARD:	PRT: 1124 AA.
AC	Q02763:		
DT	01-FEB-1994 (Rel. 28, Created)		
DT	01-FEB-1994 (Rel. 28, Last sequence update)		
DI	16-09-2003 (Rel. 43, Last annotation update)		
DE	Angiopoietin 1 receptor precursor (BC 2.7.1.112) (Tyrosine-protein kinase receptor Tib.2) (Tyrosine-protein kinase receptor TEK) (P140 TEK) (Tunica interna endothelial cell kinase).		
GN	TEK OR TIE2.		
OS	Homo sapiens (Human).		
CC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
CC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
UX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Placenta;		
RX	MEDLINE=93173509; PubMed=8382358;		
RA	Tiegey S.F., Bird T.A., Schottinger J.A., Schooley K.A., Baum P.R.;		
RT	"Molecular cloning and characterization of a novel receptor protein tyrosine kinase from human placenta.";		
RL	Oncogene 8:663-670(1993).		
RN	[2]		
RP	VARIANT VMC1 TRP-849.		
RX	MEDLINE=9714065; PubMed=8780045;		
RA	Vikula M., Boon L.M., Caraway K.L. III, Calvert T.T., Diamanti A.J.,		
RA	Goumarov B., Pasyk K.A., Marchuk D.A., Warman M.L., Cantley L.C.,		
RA	Mulliken J.B., Olse R.P.;		
RT	"Vascular dysmorphogenesis caused by an activating mutation in the receptor tyrosine kinase TIE2.";		
RL	Cell 87:1181-1190(1996).		
RN	[3]		
RP	VARIANTS VMC1 TRP-849 AND SER-897.		
RX	MEDLINE=9729245; PubMed=1053874;		
RA	Calvert T.T., Riney T.T., Kontos C.D., Cha F.H., Prieto V.G.,		
RA	Shea C.R., Berg J.N., Nevin N.C., Simpson S.A., Pasyk K.A.,		
RA	Speer M.C., Peters K.G., Marchuk D.A.;		
RT	"Allicic and locus heterogeneity in inherited venous malformations.";		
RL	Hum Mol Gene 8:1279-1289(1999)		
CC	FUNCTION: THIS PROTEIN IS A PROTEIN TYROSINE-KINASE TRANSMEMBRANE RECEPTOR FOR ANGIOPOIETIN 1. IT MAY CONSTITUTE THE EARLIEST MAMMALIAN ENDOTHelial CELL LINEAGE MARKER, PROBABLY REGULATES ENDOTHelial CELL PROLIFERATION, DIFFERENTIATION AND GUIDES THE PROPER PATTERNING OF ENDOTHelial CELLS DURING BLOOD VESSEL FORMATION.		
CC	ACTIVITY: ATP -> a protein tyrosine - ADP - protein tyrosine phosphate.		
CC	TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN ENDOTHelial CELLS AND THEIR PROGENITORS, THE ANGIOBLASTS, HAS BEEN DIRECTLY FOUND IN PLACENTA AND LUNG, WITH A LOWER LEVEL IN UMBILICAL VEIN ENDOTHelial CELLS, BRAIN AND KIDNEY.		
CC	DISSEASE: DEFECTS IN TEK ARE A CAUSE OF DOMINANTLY INHERITED VENOUS MALFORMATIONS (VMC1), AN ERROR OF VASCULAR MORPHOGENESIS CHARACTERIZED BY DILATED, SERPIGINOUS CHANNELS.		

34	179	6.9	1064	1	FBP1_STRPO	P10079 strongyloce
35	178.5	6.9	2201	1	TENA_HUMAN	P24921 homo sapien
36	178	6.8	592	1	ELL3_MOUSE	O88216 mus musculus
37	178	6.8	3110	1	LMA2_HUMAN	P24043 homo sapien
38	176	6.8	2139	1	CEB_DROME	P10040 drosophila
39	174	6.7	577	1	ITIH6_CAVPO	P18563 cavia porce
40	174	6.7	618	1	DLL3_HUMAN	O9NF77 homo sapien
41	173.5	6.7	589	1	DLL3_RAT	O88671 rattus norv
42	173.5	6.7	1376	1	CRBL_HUMAN	P82279 homo sapien
43	172.5	6.7	1154	1	LMK1_MOUSE	O61987 mus musculus
44	172.5	6.6	1798	1	LMR2_HUMAN	P55268 homo sapien
45	172	6.6	1746	1	TENA_FIC	Q25116 sus scrofa

ALIGNMENTS

CC -1- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
 CC DOMAIN.
 CC -1- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 3 FIBRONECTIN TYPE III-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.
 CC
 CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcements/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL: L06139; AAA61139.1;
 CC HSSP: P11362; 1FGK.
 CC MIM: 600221;
 CC MIM: 600195;
 CC InterPro: IPR000561; EGF-like.
 CC InterPro: IPR000719; Fok_pkinase.
 CC InterPro: IPR003961; FN-III.
 CC InterPro: IPR001245; Tyr_pkinase.
 CC Pfam: PF00008; EGF; 1.
 CC Pfam: PF00041; fn3; 3.
 CC Pfam: PF00069; pkinase; 1.
 CC PRINTS: PR00109; TYRKINASE.
 CC SMART: SM00181; EGF; 2.
 CC SMART: SM00001; EGF-like; 1
 CC SMART: SM00060; FN3; 3.
 CC SMART: SM00219; Tyrc; 1.
 CC PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 CC PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 CC PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
 CC PROSITE: PS00022; EGF_1; 3.
 CC PROSITE: PS01186; EGF_2; 3.
 CC Receptor; Tyrosine-protein kinase; Transferase, Signal, ATP-binding;
 CC Repeat; EGF-like domain; Transmembrane; Immunoglobulin domain;
 CC Glycoprotein, phosphorylation; Multigene family; Disease mutation.
 CC
 CC SIGNAL 1 18
 CC CHAIN 19 1124
 CC DOMAIN 19 745
 CC TRANSFEM 746 770
 CC DOMAIN 771 1124
 CC DOMAIN 44 102
 CC DOMAIN 210 252
 CC DOMAIN 254 299
 CC DOMAIN 301 341
 CC DOMAIN 370 424
 CC DOMAIN 444 536
 CC DOMAIN 541 634
 CC DOMAIN 638 732
 CC DOMAIN 824 1096
 CC NP_BIND 830 848
 CC BINDING 855 855
 CC ACT_SITE 964 964
 CC CARBOHYD 140 140
 CC CARBOHYD 158 158
 CC CARBOHYD 399 399
 CC CARBOHYD 438 438
 CC CARBOHYD 464 464
 CC CARBOHYD 560 560
 CC CARBOHYD 596 596
 CC CARBOHYD 649 649
 CC CARBOHYD 691 691
 CC MOD_RES 942 942
 CC VARIANT 849 849
 CC VARIANT 897 897
 CC VARIANT 1124 AA; 125810 MW; 65805018FA4QUEC CR664;
 CC SEQUENCE 1124 AA; 125810 MW; 65805018FA4QUEC CR664;
 CC
 CC Query Match 100.0% Score 2605; DR 1; Length 1124;

Best local similarity 100.0%; Pred. No. 9,70-200;
 Matches 472; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MDSLASLVLCGVSLILSGTVEGAMDLILINSLPLVSDAETSLTCLASWRPHEPTIGRD 60
 DB 1 MDSLASLVLCGVSLILSGTVEGAMDLILINSLPLVSDAETSLTCLASWRPHEPTIGRD 60
 QY 61 FEALMNOHQDPLEVTDVTFEWAQKVVWKKKASKINAYFEGEVPVKAIRIKMKMRQ 120
 DB 61 FEALMNOHQDPLEVTDVTFEWAQKVVWKKKASKINAYFEGEVPVKAIRIKMKMRQ 120
 QY 121 QASFLPATITMTVIKGINVINISPKKVIKEEDAVIYKNGSFTHSVPRHVPVLDLPHLP 180
 DB 121 QASFLPATITMTVIKGINVINISPKKVIKEEDAVIYKNGSFTHSVPRHVPVLDLPHLP 180
 QY 181 AOPDAGVYSARYIGGNLTSAFRLIVRCEAOKWGPCNHLCTACNNGVCHEDTGC 240
 DB 181 AOPDAGVYSARYIGGNLTSAFRLIVRCEAOKWGPCNHLCTACNNGVCHEDTGC 240
 QY 241 ICPGFMGRICEKACELRIFGRICNEKSCQCKSVYVPLPPYCCSCATGKGLQNE 300
 DB 241 ICPGFMGRICEKACELRIFGRICNEKSCQCKSVYVPLPPYCCSCATGKGLQNE 300
 QY 301 ACHPGFYGPCKIKRCSNNGPMCDRFGGICSPGWGICQCFREGIPRMIPKIVDLPHIE 360
 DB 301 ACHPGFYGPCKIKRCSNNGPMCDRFGGICSPGWGICQCFREGIPRMIPKIVDLPHIE 360
 QY 361 VNSGKNPICKASGWLPTNEEMTLVKPDGTVLHPDKFNHUTDHSVAITTHRLIPUSG 420
 DB 361 VNSGKNPICKASGWLPTNEEMTLVKPDGTVLHPDKFNHUTDHSVAITTHRLIPUSG 420
 QY 421 VVCSVNTVAGMVEKPFENISVKVLPKPLNAPNVDTGHNFAVINISSEFYEG 472
 DB 421 VVCSVNTVAGMVEKPFENISVKVLPKPLNAPNVDTGHNFAVINISSEFYEG 472

RESULT 2

11E2_BOVIN STANDARD: PRT: 1125 AA.
 ID TIE2_BOVIN
 AC Q06807;
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DT 16 OCT 2001 (Rel. 40, Last annotation update)
 DE Angiopoietin 1 receptor precursor (bc 2.7.1.112) (Tyrosine-protein
 DE Kinase receptor TIE-2).
 GN TEK OR TIE2 OR TIE-2.
 OS Bos taurus (Bovine).
 OC Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,
 OC Mammalia, Eutheria, Cetartiodactyla, Ruminantia, Bovidae, Bovidae,
 OC Bovidae, Bovinae; Bos.
 OX NCBI_TaxID:9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Endothelial cells;
 RX MEDLINE:9402374; PubMed:8415706;
 PA Sato T.N., Glin Y., Kozak C.A., Andus K.L.;
 FT Tie-1 and Tie-2 define another class of putative receptor tyrosine
 FT kinase genes expressed in early embryonic vascular system*;
 FL Proc. Natl. Acad. Sci. U.S.A. 90:9355-9358(1993).
 CC -1- FUNCTION: THIS PROTEIN IS A PROTEIN-TYROSINE-KINASE TRANSMEMBRANE
 CC RECEPTOR FOR ANGIOPOIETIN 1. IT MAY CONSTITUTE THE EARLIEST
 CC MAMMALIAN ENDOTHelial CELL LINEAGE MARKER. PROBABLY REGULATES
 CC ENDOTHelial CELL PROLIFERATION, DIFFERENTIATION AND GUIDES THE
 CC PROPER PATTERNING OF ENDOTHelial CELLS DURING BLOOD VESSEL
 CC FORMATION.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine -> ADP + protein
 CC tyrosine phosphate.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- TISSUE SPECIFICITY: SPECIFICALLY EXPRESSED IN DEVELOPING VASCULAR
 CC ENDOTHelial CELLS.
 CC -1- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
 CC DOMAIN.
 CC -1- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.

CC -1- SIMILARITY: CONTAINS 3 FIBRONECTIN TYPE III-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.
 CC This SWISS-PROT entry is copyright, it is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@sib-sib.ch).
 CC -----

DR EMBL: X71434; CAAS0555.1; -
 DR PIR: S42691; S32691.
 DR HSSP: P11362; IFGK.
 DR InterPro: IPR003561; EGF-like.
 DR InterPro: IPR007019; Euk pk kinase.
 DR InterPro: IPR003961; FN III
 DR InterPro: IPR002049; Laminin_EGF.
 DR InterPro: IPR001245; Tyr_kinase.
 DR Pfam: PF00008; EGF_2.
 DR Pfam: PF00041; fn3; 3.
 DR Pfam: PF00069; pk kinase; 1.
 DR PRINTS: PR00109; IYRKINASE.
 DR SMART: SM00180; EGF_Lam.
 DR SMART: SM00001; EGF_Like; 1.
 DR SMART: SM00060; FN3; 3.
 DR SMART: SM00219; TyrKc; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_AIP; 1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00022; EGF_1; 3.
 DR PROSITE: PS01186; EGF_2; 2.
 KW receptor, tyrosine protein kinase, transferase, signal, ATP-binding;
 KW repeat, EGF-like domain, transmembrane, immunoglobulin domain,
 KW glycoprotein; Phosphorylation; Multigene family.
 FT SIGNAL 1 18
 FT CHAIN 19 1125
 FT DOMAIN 19 746
 FT TRANSMEM 747 771
 FT POTENTIAL.
 FT CYTOPLASMIC (POTENTIAL).
 FT IGF-LIKE C2-TYPE DOMAIN 1
 FT EGF-LIKE 1.
 FT EGF-TYPE 2.
 FT EGF-LIKE 3.
 FT IGF-LIKE C2-TYPE DOMAIN 2.
 FT FIBRONECTIN TYPE-III 1.
 FT FIBRONECTIN TYPE-III 2.
 FT FIBRONECTIN TYPE-III 3.
 FT PROTEIN KINASE.
 FT NP_BIND 831 839
 FT ATP (BY SIMILARITY).
 FT ATP (BY SIMILARITY).
 FT HY SIMILARITY.
 FT PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT N-LINKED (GLCNAC...) (POTENTIAL)
 FT CARBOHYD 158 158
 FT SEQUENCE 1125 AA; 125927 MW; 015F1320AB853B7F CRC64;

Query: March
 Best local similarity 93.5%; Score 2435, DR 1, Length 1125,
 Matches 440; Conservative 12; Mismatches 20; Indels 0; Gaps 0;
 QY 1 MDSLASLVLCGVSLLSGTGVSGAMDLINSLPLVSDAETSLTCTASWPPHEPITIGPD 60
 DB 1 MDSLAGVLCGVSLLSAIVDGMADLILINSLPLVSDAETSLTCTASWPPHEPITIGPD 60
 QY 61 FEALMCHQDPLEVTADVTPEWAKKVVWPERKASKINAYPFGVPVPEAIPITMKMPQ 120
 DB 61 FEALMCHQDPLEVTQDATPEWAKKVVWPERKASKINAYFGVPEVPEAIPITMKMKQ 120
 QY 121 QASFLPATITMTVDKGVNVSFKKVLKPEDAVLYKNSFIRHSPRHEVVDILLEVHLP 180
 DB 121 QASFLPATITMTVDKGVNVSFKKVLKPEDAVLYKNSFIRHSPRHEVVDILLEVHLP 180

QY 181 AQPDAQWYSAPYIIGNIFTSAFELIVPRFEAKWKPFENHIDATAMNNVWVHEITSEC 240
 DB 181 AQPDAQWYSAPYIIGNIFTSAFELIVPRFEAKWKPFENHIDATAMNNVWVHEITSEC 240
 QY 241 ICPPHFMKTEKACAEHLHFKICKERCSQDQCKSVFCLPDPYQSCAIGWKGACQNE 300
 DB 241 ICPPHFMKTEKACAEHLHFKICKERCSQDQCKSVFCLPDPYQSCAIGWKGACQNE 300
 QY 301 AFHAEKYSPLAKIKCSNNNHMTAFQWGLCSQWGLQCEKPEGLPQWMLPKVDPHPHIF 360
 DB 301 AFHAEKYSPLAKIKCSNNNHMTAFQWGLCSQWGLQCEKPEGLPQWMLPKVDPHPHIF 360
 QY 361 VNSKFPNDPKASQWPLTBEEMTLVKPDTVLHFKPKNHTDMPFSAITHTILPTDSG 420
 DB 361 VNSKFPNDPKASQWPLTBEEMTLVKPDTVLHFKPKNHTDMPFSAITHTILPTDSG 420
 QY 421 VVCSVNTVACMVEKPNISVKVLKPEINAFNVDTGCHFAV:NISSEPVFG 472
 DB 421 VVCSVNTVACMVEKPNISVKVLKPEINAFNVDTGCHFAV:NISSEPVFG 472
 RESULT 3
 TIE2_MOUSE
 ID TIE2_MOUSE STANDARD: PRT: 1122 AA.
 AC Q02858;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1995 (Rel. 31, last sequence update)
 DE 16-OCT-2001 (Rel. 40, last annotation update)
 DE Argiope toxin 1 receptor precursor (ic 2.7.1.12) (Tyrosine-protein
 DE kinase receptor TIE-2) (Tyrosine-protein kinase receptor TEK) (P140
 DE TEK) (Tunica interna endothelial cell kinase) (HYK).
 DE TEK or TIE2 or TIE-2 or HYK
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Molecostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BAIRYC; TISSUE=Lung;
 RA MEDLINE=94022374; PubMed=8415706;
 FT Sato T.N., Qin Y., Kozak G.A., Andus K.B.,
 FT "Tie-1 and tie-2 define another class of putative receptor tyrosine
 FT kinase genes expressed in early embryonic vascular system.";
 FT Proc Natl Acad Sci U S A 90:9355-9358(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CO-1; TISSUE=Embryonic heart;
 RA MEDLINE=93241731; PubMed=8386827;
 FT Dumont D.T., Gradwohl G.T., Foaq G.H., Auerbach R., Breitman M.L.;
 FT "The endothelial-specific receptor tyrosine kinase, tek, is a member
 FT of a new subfamily of receptors.";
 FT Oncogene 8:1293-1301(1993).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Embryonic stem cells;
 RA MEDLINE=93129253; PubMed=1283811;
 FT Horita K., Yagi T., Kohmura N., Tomooka Y., Ikawa Y., Aizawa S.;
 FT "A novel tyrosine kinase, hyk, expressed in murine embryonic stem
 FT cells.";
 FT Biochem Biophys Res Commun 189:1747-1753(1992).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RA MEDLINE=94031116; PubMed=8217221;
 FT Kuntz A.S., Stacker S.A., Wilks A.P.;
 FT "Tie2, a putative protein tyrosine kinase from a new class of cell
 FT surface receptor.";
 FT Growth Factors 9:99-105(1993).
 RN [5]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=94244487; PubMed=8187650;
 FT Schuerch H., Risau W.;
 FT "Expression of tie-2, a member of a novel family of receptor tyrosine

kinases, in the endothelial cell lineage.*;
 [6]
 SEQUENCE OF 822-1122 FROM N.A.
 STRAIN-CD-1; TISSUE-Embryonic heart;
 MEDLINE-92434855; PubMed-1630810.
 RA Dumont D.J., Yamauchi T.P., Coulon R.A., Rossant J., Breitman M.L.;
 *Tek, a novel tyrosine kinase gene located on mouse chromosome 4, is
 expressed in endothelial cells and their presumptive precursors.*;
 RL oncogene 7:1471-1480(1992).
 CC -1- FUNCTION: THIS PROTEIN IS A PROTEIN TYROSINE-KINASE TRANSMEMBRANE
 RECEPTOR FOR ANGIOPOIETIN 1. IT MAY CONSTITUTE THE EARLIEST
 MAMMALIAN ENDOTHELIAL CELL LINEAGE MARKER. PROBABLY REGULATES
 ENDOTHELIAL CELL PROLIFERATION, DIFFERENTIATION AND SHIELDS THE
 PROPER PATTERNING OF ENDOTHELIAL CELLS DURING BLOOD VESSEL
 FORMATION.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein-tyrosine- -> ADP + protein
 tyrosine phosphate.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- TISSUE SPECIFICITY: SPECIFICALLY EXPRESSED IN DEVELOPING VASCULAR
 ENDOTHELIAL CELLS.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSION DETECTABLE IN DAY 8.5 EMBRYOS
 CC -1- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
 DOMAIN.
 CC -1- SIMILARITY: CONTAINS 2 IMMUNOGLULIN-LIKE C2-TYPE DOMAINS
 CC -1- SIMILARITY: CONTAINS 3 FIBRONECTIN TYPE III-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.
 CC
 CC This Swiss-prot entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/annotation/>
 or send an email to license@sib-sib.ch).
 CC
 EMBL: X71426; CAA50557.1;
 EMBL: X67553; CAA42857.1;
 EMBL: D13738; BAA02883.1;
 EMBL: S67051; AAB28663.1;
 PIR: S33142; S33142.
 HSSP: P11362; IFGK.
 MGI: MGI:98664; Tek.
 InterPro: IPR000561; EGF-like
 InterPro: IPR000719; Euk_kinase.
 InterPro: IPR003961; FNIII.
 InterPro: IPR001245; Tyr_kinase.
 Pfam: PF00041; In3; 3.
 Pfam: PF00069; pkinase; 1.
 PRINTS: PR00109; TYRKINASE.
 SMART: SM00181; EGF; 2.
 SMART: SM00069; FN3; 2.
 SMART: SM00219; Tyrc; 1.
 PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.
 PROSITE: PS00122; EGF_1; 3.
 PROSITE: PS01186; EGF_2; 3.
 Receptor: Tyrosine protein kinase; Transferase, Signal; APP-binding;
 KW Repeat: EGF-like domain; Transmembrane, Immunoglobulin domain;
 KW Glycoprotein; Phosphorylation; Multigene family.
 FT SIGNAL 1 18 POTENTIAL.
 FT CHAIN 19 1122 ANGIOPOIETIN 1 RECEPTOR.
 FT DOMAIN 19 744 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 745 769 POTENTIAL.
 FT DOMAIN 779 1122 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 44 102 IGF-LIKE C2-TYPE DOMAIN 1.
 FT DOMAIN 219 252 EGF-LIKE 1.
 FT DOMAIN 254 299 EGF-LIKE 2.
 FT DOMAIN 301 341 EGF-LIKE 3.
 FT DOMAIN 370 424 IGF-LIKE C2-TYPE DOMAIN 2.
 FT DOMAIN 444 536 FIBRONECTIN TYPE-III 1
 FT DOMAIN 539 633 FIBRONECTIN TYPE-III 2.

FT DOMAIN 637 731 FIBRONECTIN TYPE-III 3.
 FT DOMAIN 822 1094 PROTEIN KINASE
 FT NP_BINDING 828 836 AIP (BY SIMILARITY).
 FT BINDING 853 853 AIP (BY SIMILARITY).
 FT ACT_SITE 962 962 BY SIMILARITY.
 FT CARBOHYD 146 146 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 158 158 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 359 359 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 438 438 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 464 464 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 558 558 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 595 595 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 648 648 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 690 690 N-LINKED (GLCNAC...) (POTENTIAL).
 FT MOD_RES 990 990 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT CONFLICT 161 171 FHSVPRHVP -> LHPLSAUGMKYL (IN REF. 3).
 FT CONFLICT 538 538 S -> C (IN REF. 1 AND 4).
 FT CONFLICT 736 736 A -> G (IN REF. 1 AND 4).
 FT CONFLICT 745 761 MLLIALGSGATCTIV -> DATSHHPVWVWDFASPC
 (IN REF. 3).
 FT CONFLICT 786 786 N -> NV (IN REF. 3).
 FT CONFLICT 913 913 R -> G (IN REF. 3).
 FT CONFLICT 925 931 ALANSTA -> CHROOYS (IN REF. 3).
 FT CONFLICT 1117 1117 S -> P (IN REF. 3).
 SQ SEQUENCE 1122 AA: 1257aa MW: 187962.30 Da pI: 4.96 CRC64:
 Query Match 92.4%; Score 2407; DB 1; Length 1122;
 Best Local Similarity 91.9%; Pred. No. 5,70-184;
 Matches 434; Conservative 17; Mismatches 21; Indels 0; Gaps 0;
 QY 1 MDSLASLVAGVSHLSGIVVERAMGHTILNSTPLVDAELISLTICIASGRPHIEPTIGRD 60
 DB 1 MSLIAGVLCGVSHIYGVVVGAMUILLNSLPLVSVAPISLTICIASGRPHIEPTIGRD 60
 QY 61 FEALMNGHQDPLEVTQDVTREWAKKVVWKEKASKINGAYFCEGSPVGEAIPRTMKMPQ 120
 DB 61 FEALMNGHQDPLEVTQDVTREWAKKVVWKEKASKINGAYFCEGSPVGEAIPRTMKMPQ 120
 QY 121 QASPLPATLIMTVKGDVNISEKVKLIKEDDAVIYKNGSFTHSVPRHVPDILEVHLPH 180
 DB 121 QASPLPATLIMTVKGDVNISEKVKLIKEDDAVIYKNGSFTHSVPRHVPDILEVHLPH 180
 QY 181 AQPDACVYVYAPYVGNLFTSAPTRIVPPCEACKWGPEDCNELCATCMNNVYCHETGEV 240
 DB 181 AQPDACVYVYAPYVGNLFTSAPTRIVPPCEACKWGPEDCNELCATCMNNVYCHETGEV 240
 QY 241 ICPPGEMGRTCEKACELHTGRTCKERCQSQBCKSVYVFLPDPYGSCTATGKGLQNE 300
 DB 241 ICPPGEMGRTCEKACELHTGRTCKERCQSQBCKSVYVFLPDPYGSCTATGKGLQNE 300
 QY 301 ACHRGVYGPWTKIACSCNNGEMCTHAFQGNICSPGWQGLCCPRPGIIPRMTPKIVDLPDHF 360
 DB 301 ACHRGVYGPWTKIACSCNNGEMCTHAFQGNICSPGWQGLCCPRPGIIPRMTPKIVDLPDHF 360
 QY 361 VNSCKFNPTCKASQWPLPTNFEMTLVPPCTVLHPKDFNHLHESVAISFHLHLLPDSG 420
 DB 361 VNSCKFNPTCKASQWPLPTNFEMTLVPPCTVLHPKDFNHLHESVAISFHLHLLPDSG 420
 QY 421 VWVSVNTVAGVMEKPEPNISVKVLPKPLNAPNVDTGHNFAVINISSEPYFG 472
 DB 421 VWVSVNTVAGVMEKPEPNISVKVLPKPLNAPNVDTGHNFAVINISSEPYFG 472
 RESULT 4
 TIE1_BOVIN
 ID TIE1_BOVIN STANDARD: PRT: 1136 AA.
 AC Q06805,
 DT 01-FEB-1995 (Rel. 31, Created)
 DT 01-FEB-1995 (Rel. 41, Last Sequence update)
 DT 15-JUL-1999 (Rel. 48, Last annotation update)
 DE Tyrosine-protein kinase receptor TIE-1 precursor (90.2.7.1.112).
 GN TIE1 OR TIE-1.

OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Endothelial cells;
 RA MEDLINE=94022374, PubMed=8415706,
 PA Sato T N., Qin Y., Kozak C A., Andus K L.
 RT "Tie-1 and tie-2 define another class of putative receptor tyrosine
 kinase genes expressed in early embryonic vascular system.";
 RI Proc. Natl. Acad. Sci. U.S.A. 90:9355-9358(1993).
 CC -!- FUNCTION: PREPARE PROTEIN TYROSINE KINASE TRANSMEMBRANE RECEPTOR.
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- TISSUE SPECIFICITY: SPECIFICALLY EXPRESSED IN DEVELOPING VASCULAR
 CC ENDOTHELIAL CELLS.
 CC -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
 CC DOMAIN.
 CC -!- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC -!- SIMILARITY: CONTAINS 3 FIBRONECTIN TYPE III-LIKE DOMAINS.
 CC -!- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 DR EMBL: X71423; CAA50554.1; -
 DR PIR: S32690; S32690.
 DR HSP: P11362; LFQK.
 DR InterPro: IPR0003561; EGF like.
 DR InterPro: IPR000719; Euk_FKinase.
 DR InterPro: IPR003961; FN_III.
 DR InterPro: IPR003599; Iq.
 DR InterPro: IPR003006; Iq_MHC.
 DR InterPro: IPR001245; Tyr_FKinase.
 DR Pfam: PF00008; EGF; 2.
 DR Pfam: PF00041; fn3; 3.
 DR Pfam: PF00047; Iq; 2.
 DR Pfam: PF00069; pKinase; 1.
 DR PRINTS: PR00109; TYRKINASE.
 DR SMART: SM00181; EGF; 2.
 DR SMART: SM00060; FN3; 2.
 DR SMART: SM00409; Iq; 1.
 DR SMART: SM00219; TyrcK; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS00022; EGF_1; 3.
 DR PROSITE: PS01186; EGF_2; 3.
 KW Receptor; Tyrosine protein kinase; Transferase; Signal; AIP-binding;
 KW Repeat; EGF-like domain; Transmembrane; Immunoglobulin domain;
 KW Glycoprotein; Phosphorylation; Multigene family.
 FT SIGNAL: 1 23
 FT CHAIN: 24 1136
 FT DOMAIN: 24 757
 FT EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM: 758 782
 FT DOMAIN: 783 1136
 FT CYTOPLASMIC (POTENTIAL)
 FT DOMAIN: 43 106
 FT Iq-LIKE C2-TYPE DOMAIN 1.
 FT EGF LIKE 1.
 FT DOMAIN: 212 254
 FT EGF LIKE 2.
 FT DOMAIN: 256 301
 FT EGF LIKE 3.
 FT DOMAIN: 303 343
 FT Iq-LIKE C2-TYPE DOMAIN 2.
 FT DOMAIN: 370 424
 FT FIBRONECTIN TYPE-III 1.
 FT DOMAIN: 444 538
 FT FIBRONECTIN TYPE-III 2.
 FT DOMAIN: 541 637
 FT FIBRONECTIN TYPE-III 3.
 FT DOMAIN: 641 742
 FT PROTEIN KINASE.
 FT DOMAIN: 837 1116

FT NP_BIND 843 851 ATP (BY SIMILARITY).
 FT BINDING 858 868 AIP (BY SIMILARITY).
 FT ACT_SITE 977 977 BY SIMILARITY.
 FT MOD_RES 1005 1005 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT CARBOHYD 84 84 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 159 159 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 501 501 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 594 594 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 707 707 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 1136 AA: 124953 MW: 117580.434 kDa
 Query Match 33.58, Score 873.5, DB 1, Length 1136;
 Best Local Similarity 33.38, Pred. No. 8,20,62;
 Matches 179; Conservative 68; Mismatches 176; Indels 33; Gaps 12;
 QY 16 LSGTVEGAMDLILNSLFLVSDAETSLTCL -- ASGWRPHEPIITIGRDFEALMN 66
 DB 15 LASHGGAAYDITLLADIRLTPQRFFLTVCSEACAGAGSGDAWGP--PLLEKKDRIVRT 72
 QY 67 QR--QDPLEVTQVTPREWAKKVKMKP-EKASKINAYEPFEPVPEAEIPIPTMKMRQAS 123
 DB 73 PRPWQPP-----HIAPNGSSPTVTRGFSQPSDLVCFVSCVG---GGGTPELVYVHNSGAI 124
 QY 124 PLPATLITMDVKGDVNIISPKKVLIKEDDAVIYKNGSFHSVPREVID---ILEVHLPH 180
 DB 125 ILPOKVTHTVANKGDAVLASARVPKPLTDVITWKNSSYEYTLDEHAGDQVFLQ--LPN 182
 QY 181 APLPLAIAWYSAPYILGSLFTSAFTPLIVPRPEAKWKPE-NHLITAMNNWVHEDFEE 240
 DB 183 VQPSGGIYSAIYLEASPLSSAFTPLIVPQFAKPPWQDLIKEDQDLHGVCHQNGRC 242
 QY 241 ICPPEFMGPTTEKATEIHTPFTPEKPSGAGGKSVVFLPDPYSGSCATGKWLQONE 300
 DB 243 VCPKFTTTRTEALVPEGRFASQVGEQEPITSGTPGLTFPLKPEVSSSSSWPSSQGE 302
 QY 301 ACHHTPTPTPTKCTKCTSNKNTNNTNNTNNTNNTNNTNNTNNTNNTNNTNNTNNTNNT 350
 DB 303 ACAPGFFAN"ILG"Q"UN"NTNNTNNTNNTNNTNNTNNTNNTNNTNNTNNTNNTNNTNNT 359
 QY 361 VNSGKFNPI-CKASGMPLENTLVKPDGTV-IHPKDFNHTDHFSAVFTIIRLPPD 418
 DB 360 FNIDIMPINCAAGANPFPVWQSMFLKAPPDIVLISUKAVFEDW-LIAFEPVPLAIGD 418
 QY 419 SCVWVCVNTVAGVVEKPFNISVVKPLPKPLNAPVI 454
 DB 419 SCLWEPVSTSGQDSRRFRINVKVPPVPLTAPRL 454
 RESULT 5
 TIE1_HUMAN
 ID TIE1_HUMAN STANDARD: PRT: 1138 AA.
 AC F35590;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-JUN-1999 (Rel. 38, Last annotation update)
 DE Tyrosine protein kinase receptor TIE-1 precursor (EC 2.7.1.112).
 GN TIE1 OR TIE.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP MEDLINE=94195316; PubMed=1312667;
 RA Fattanen J., Armstrong E., Mackelae T.P., Korhonen J., Sandberg M.,
 RA Reinkonen K., Knuutila S., Huhtani K., Allitalo K.;
 RT "A novel endothelial cell surface receptor tyrosine kinase with
 RT extracellular epidermal growth factor homology/ domains.";
 RL Mol. Cell. Biol. 12:1698-1707(1992).
 RN [2]
 PP REVISIONS.
 RA Fattanen J.M.;

Submitted (JUL 1993) to the EMBL/GenBank/GenBank databases
 CC 1- FUNCTION: PROBABLE PROTEIN TYROSINE KINASE TRANSMEMBRANE RECEPTOR.
 CC 1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein
 CC tyrosine phosphate.
 CC 1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC 1- TISSUE SPECIFICITY: SPECIFICALLY EXPRESSED IN DEVELOPING VASCULAR
 CC ENDOTHELIAL CELLS.
 CC 1- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
 CC DOMAIN.
 CC 1- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC 1- SIMILARITY: CONTAINS 3 FIBRONECTIN TYPE III-LIKE DOMAINS.
 CC 1- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.
 CC
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb.sib.ch/announce/>
 CC or send an email to license@sib-sb.ch).

EMBL: X60957; CAA43290.1; -

HSSP: P11462; IFGK.

MIM: 600222; -

InterPro: IPR000561; EGF-like.

InterPro: IPR000719; Euk_pkinase.

InterPro: IPR000661; FN_III.

InterPro: IPR004006; Iq_MHC.

InterPro: IPR001245; Tyr_pkinase.

Pfam: PF00008; EGF_2.

Pfam: PF00041; fn3_3.

Pfam: PF00047; Ig_2.

Pfam: PF00069; pkinase; 1.

PRINTS: PR00109; TYRKINASE.

SMART: SM00181; EGF_2.

SMART: SM00060; FN3_2.

SMART: SM00219; TyPKc_1.

PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.

PROSITE: PS00100; PROTEIN_KINASE_TYR; 1.

PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.

PROSITE: PS00022; EGF_1; 3.

PROSITE: PS01186; EGF_2; 3.

RECEPTOR_TYROSINE-protein kinase; Transferase, Signal, ATP-binding;

Repeat: EGF-like domain; Transmembrane; Immunoglobulin domain;

Glycoprotein; Phosphorylation; Multigene family.

Signal 1 24 POTENTIAL.

CHAIN 25 1138 TYROSINE-PROTEIN KINASE RECEPTOR TIE-1.

DOMAIN 25 759 EXTRACELLULAR (POTENTIAL).

TRANSMEM 760 784 POTENTIAL.

DOMAIN 785 1138 CYTOPLASMIC (POTENTIAL).

DOMAIN 43 105 IGF-LIKE C2-TYPE DOMAIN 1.

DOMAIN 214 256 EGF-CTF 1.

DOMAIN 258 303 EGF-LIKE 2.

DOMAIN 305 345 EGF-LIKE 3.

DOMAIN 372 426 IGF-LIKE C2-TYPE DOMAIN 2.

DOMAIN 446 540 FIBRONECTIN TYPE-III 1.

DOMAIN 543 639 FIBRONECTIN TYPE-III 2.

DOMAIN 643 744 FIBRONECTIN TYPE-III 3.

DOMAIN 839 1118 PROTEIN KINASE.

NP-HIND 845 853 ATP (HY SIMILARITY).

HINDING 870 870 ATP (HY SIMILARITY).

ACT SITE 979 979 HY SIMILARITY.

CARBOHYD 83 83 N-LINKED (GLCNAC...) (POTENTIAL).

CARBOHYD 161 161 N-LINKED (GLCNAC...) (POTENTIAL).

CARBOHYD 503 564 N-LINKED (GLCNAC...) (POTENTIAL).

CARBOHYD 596 596 N-LINKED (GLCNAC...) (POTENTIAL).

CARBOHYD 709 709 N-LINKED (GLCNAC...) (POTENTIAL).

MOD_RES 1067 1007 PHOSPHORYLATION (AUTC...) (HY SIMILARITY).

SEQUENCE 1138 AA: 125389 MW: 38428336760581 CRO64.

Query Match

Best local similarity 38.2%; Score 865.5; DB 1; Length 1138;

Pred. No. 3.6e-61;

Matches 174, Conservative 71, Mismatches 185, Indels 25, Gaps 8;
 QY 14 LLLSTVEGAMLLILLINSLPLVSDAETSLTCL-----ASWEPHEPITIGRDFEAL 64
 DB 13 LFLASHWAADVLLLANLKLIDLPQFELICVSSAAGAGSGDAWGP--PLLEKIDRIV 70
 QY 65 MNOHQDPLEVTDQVIREWAKKVVWKR--EKASKINGAVFQFGRVGRGHAIRIMKMHQOAS 123
 DB 71 RTPGPELR---LARNCSHQVTLIPAFSPKSTDLGVFSVSGAGAPTPVIVYINSFGAH 126
 QY 124 FLPATLMTVDKNDVNIISFKVLLIKEEDAVIYKNSIFHSVPRHEVDP--ILEVHLPH 180
 DB 127 LLDPKVTHTVNGKQIAVLSARVHKQKIWIKNSYFYTLLWHEAQKRFILQ--LPN 184
 QY 191 AGLGAGVAFVAFYTCGLNFTSAFTPIVPPCEAGKNGCEPCHLCTACMNNWCHETPEC 240
 DB 185 VQPESSNYIYNTLIEASPLSAFTFLIVRGAGVHWHGATIKPAGGLDQVYTHIHDEE 244
 QY 241 TQPPPEMERICEKACPIHTFERCKEPCSPQEKSVVFEIPIQYQCSFALGKGLQNE 300
 DB 245 VCPPEFTTTCLEAGREGREFSQSQDQTPGLISGRGLLPLLPYQCSGSGMWSQVQE 304
 QY 361 ACHFAFVYFPAKIPSSNNSEMTPKFGGLSPGWGLQVCEPEEIPMTPIVDLPDITE 360
 DB 305 ACAPGHFGADCRLOCOCONGTCDRFSGCVPSGWHGVCHCKSD---RIPGLNMASELE 363
 QY 361 VNSCKENFI--CKASGWPLFTNEEMTLVKFGDGVTHIEKDNHTDHSVAIFTIIRLPDS 419
 DB 362 FNLETMPPIVNCAGNPPFVPGSGSTELPKPQCTVLI--STKAIPEPEKTTAEFEVPLVADS 421
 QY 420 GVWCVSVNTVAGVVEKPFNISVKVLDPKPLNAPNVI 454
 DB 422 GFWEFRVSTSGGATSPSPKPKVNVKVPVPLAARLL 456

RESULT 6

TIE1_MOUSE

ID TIE1_MOUSE STANDARD: PRT: 1134 AA.

AC Q06806;

DT 01-FEB-1995 (Rel. 31, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Tyrosine protein kinase receptor TIE-1 precursor (NC 2.7.3.112).

GN TIE1 OR TIE-1.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID:10090;

RN 11

RP SEQUENCE FROM N.A.

RC STRAIN-BALB/C; TISSUE-Lung;

PX MEDLINE:94022374; PubMed-8415706;

PA Sato T N., Qin Y., Kovak C A., Andrus K L.;

RT "Tie 1 and tie-2 define another class of putative receptor tyrosine

kinase genes expressed in early embryonic vascular system.";

RL Proc. Natl. Acad. Sci. U.S.A. 90:9355-9358(1993).

RN 12

RP SEQUENCE FROM N.A.

RC STRAIN-BALB/C; TISSUE-Bone marrow;

RA Krivosov A.V., Frisler M.A., Visser J.W.M., Helyavsky A.V.;

RL Submitted (JUN-1994) to the EMBL/GenBank/DBJ databases.

RN 13

RP SEQUENCE FROM N.A.

RC STRAIN BALB/C; TISSUE-Liver;

PX MEDLINE:9331441; PubMed-8395848;

PA Iwama A., Hamauchi I., Hashiyama M., Murayama Y., Yasunaga K.,

RA Suda T.;

RT "Molecular cloning and characterization of mouse TIE and TEK receptor

tyrosine kinase genes and their expression in hematopoietic stem

cells.";

RL Biochem. Biophys. Res. Commun. 195:301-309(1993).

RN 14

RP SEQUENCE OF 1-19 FROM N.A.

RC STRAIN=129/SV, TISSUE: Liver.
 RA MEDLINE=95387653; PubMed=7655012;
 RX Korhonen J., Laitinen I., Halmekyto M., Alhonen L., Janne J.,
 RA Dumont D., Allitalo K.:
 RA "Endothelial-specific gene expression directed by the tie gene
 RT promoter in vivo.";
 PL Blood 86:1828-1836(1995).
 CC -!- FUNCTION: PROBABLY PROTEIN TYROSINE KINASE TRANSMEMBRANE PROTEIN
 CC -!- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein
 CC tyrosine phosphate
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- TISSUE SPECIFICITY: SPECIFICALLY EXPRESSED IN DEVELOPING VASCULAR
 CC ENDOTHELIAL CELLS.
 CC -!- SIMILARITY: TO OTHER PROTEIN-TYROSINE KINASES IN THE CATALYTIC
 CC DOMAIN.
 CC -!- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
 CC -!- SIMILARITY: CONTAINS 3 FIBRONECTIN TYPE III-LIKE DOMAINS.
 CC -!- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.
 CC This SWISS-PROT entry is copyright, it is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcements/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: X71425; CAA50556.1; -;
 DR EMBL: X80764; CAA56736.1; -;
 DR EMBL: X73960; CAA52148.1; -;
 DR EMBL: S79146; -; NOT_ANNOTATED; CDS
 DR HSPB: P11362; IFGK
 DR MGD: M51-64466; T1-1
 DR InterPro: IPR000561; EGF-like.
 DR InterPro: IPR000719; Euk_Pkinase
 DR InterPro: IPR003961; FN_III.
 DR InterPro: IPR003599; Iq.
 DR InterPro: IPR003006; Iq_MHC.
 DR InterPro: IPR004600; Iq_Like
 DR InterPro: IPR001245; Tyr_Pkinase.
 DR Pfam: PF00008; EGF; 2.
 DR Pfam: PF00041; fn3; 3.
 DR Pfam: PF00047; Iq; 2.
 DR Pfam: PF00069; pkinase; 1.
 DR PRINTS: PS00109; TYRKINASE
 DR SMART: SM00181; EGF; 2.
 DR SMART: SM00060; FN3; 2.
 DR SMART: SM00022; EGF_1; 1.
 DR SMART: SM00409; IG; 1.
 DR SMART: SM00219; IG_Like; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_Atp; 1.
 DR PROSITE: PS00109; PROTEIN_KINASE_Tyr; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1
 DR PROSITE: PS00022; EGF_1; 1
 DR PROSITE: PS01186; EGF_2; 3.
 DR Receptor_Tyrosine_Protein_Kinase_Transmembrane_Immunoglobulin_domain.
 KW Repeat: EGF-like domain; Transmembrane; Immunoglobulin domain.
 KW Glycoprotein, Phosphorylation; Multi-subunit family.
 FT SIGNAL 1 22
 FT CHAIN 23 1134 TYROSINE PROTEIN KINASE RECEPTOR TIE 1.
 FT DOMAIN 23 755 EXTRACELLULAR (POTENTIAL).
 FT TRANSFEM 756 780 POTENTIAL.
 FT DOMAIN 781 1134 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 43 103 IG-LIKE C2-TYPE DOMAIN 1.
 FT DOMAIN 212 254 EGF-LIKE 1.
 FT DOMAIN 256 301 EGF-LIKE 2.
 FT DOMAIN 303 343 EGF-LIKE 3.
 FT DOMAIN 370 424 IG-LIKE C2-TYPE DOMAIN 2.
 FT DOMAIN 444 538 FIBRONECTIN TYPE-III 1.
 FT DOMAIN 541 635 FIBRONECTIN TYPE-III 2.
 FT DOMAIN 639 740 FIBRONECTIN TYPE-III 3.
 FT DOMAIN 835 1114 PROTEIN KINASE.
 FT NO_BIND 841 849 ATP (BY SIMILARITY).

PI BINDING 806 806 AIP (BY SIMILARITY).
 FT ACT_SITE 975 975 BY SIMILARITY.
 FT MOD_RES 1003 1003 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT CARBOHYD 81 81 N-LINKED (GLCNAC) (POTENTIAL).
 FT CARBOHYD 159 159 N-LINKED (GLCNAC) (POTENTIAL).
 FT CARBOHYD 501 501 N-LINKED (GLCNAC) (POTENTIAL).
 FT CARBOHYD 547 547 N-LINKED (GLCNAC) (POTENTIAL).
 FT CARBOHYD 705 705 N-LINKED (GLCNAC) (POTENTIAL).
 FT CONFLICT 599 599 R -> L (IN REF. 1).
 SQ SEQUENCE 1134 AA; 1246 aa MW: 66126 Da; pI: 5.6; PK64;

Query Match 31.5%; Score 820.5, DB 1, Length 1134;
 Host local Similarity 36.0%; Pred. No. 146-57;
 Matches 169; Conservative 70; Mismatches 191; Indels 39; Gaps 7;

QY 5 ASLVGVGSLLSGVGAMDLILNLSPLVSDAETSLTICIA---SGWRPHEPILIGRD 60
 DB 6 SSLLD--PTLFLASHVGASVDLTLLANLKTIDPQRFHICVSGACAGRSSDPLLILK 63
 QY 61 -----FEALMNHQHDPLEVTVQVTFHWAKKVVWREKASKINAYPTPRVP 107
 DB 64 LRTVPTFPQQLYLAAPNASH-VILP-----GESKPSLLVGVESHVWAG 108
 QY 108 GRAPIPIKMKMGLASFPAILIMIVKGVNUNISPKKVLKEKDAVYKNSPFIHSVP 167
 DB 109 ARPTRVLYVNSPGAHFPKPKVTHTVNKGDVTLVAHVHKETQVIMKNGSYENTLDW 168
 QY 168 HEVPO-ILEVHLPDAQPJGAVSYAYIGSLFTSAFTPLIVPPTEAKWKEPENILCTA 226
 DB 169 QFADNCPQVQIQNVQPSLSIYSAIYFASQCSAPPPIVPGCCAGPWGCGVKKIC 228
 QY 227 QNNNVWRELEETTFPMSTEFKVELHTTETVEFSSGEPFSPVYVLEFETYG 286
 DB 229 TLFWVWREHGFQVFTFTTETFEQAPEGFQSGSGVEFETSTAYFVLLLEFETYG 288
 QY 287 QSCATGKGLGQNEACHPGFYGQCKLKCSNNHGMCKDFQSCIASPGWGLCKERGLP 346
 DB 289 QSGSGWPKSGSLEAAFGRFJADQLYQLVNGSTCFSSGVTSQWGWGKESD-- 346
 QY 347 RMTPKVIDLPDHEVNSGKNEPI-CKASGWLPTNEEMTLVKPDGTVLHPKDNITDIHS 405
 DB 347 -KIQQLISMAITVFNIGIMPIVNCAAAQNPFPVQSMALPRPGLMILSLKIVIVPDKT 405
 QY 406 VALFTIHPILPDSGVWVSVNTVAGWKEKPENISVKKVLPKPLNAPNVI 454
 DB 406 TAEFEVPSLTLHSGSGFWETKYSTSGSQSGSPREKPVNVKVPVPTTAPPELL 454

RESULT 7
 SPEC_HUMAN
 ID SPEC_HUMAN STANDARD: PPT: 830 AA.
 AC Q14152; Q43701;
 DT 01 MAR 2002 (Ref. 41, Created)
 DI 01 MAR 2002 (Ref. 41, Last sequence update)
 DI 01 MAR 2002 (Ref. 41, Last annotation update)
 DE Endothelial cellis scavenger receptor precursor (Acr-1); LIL (non-polar).
 GN SPC10P KIAA0149
 OS Homo sapiens (Human).
 CC Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Mammalia, Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Umbilical vein endothelial cells;
 RX MEDLINE=98058897; PubMed=9395444;
 RA Adachi H., Tsujimoto M., Arai H., Inoue K.:
 RT "Expression cloning of a novel scavenger receptor from human
 endothelial cells."
 PL J. Biol. Chem. 272:31217-31220(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Bone marrow;

Db 385 YEGG-----VKNCPENTGWNKNSGYPPRPVWNNHQCKVCQDDPKHT 435
 Qy 231 GVCHEDTCTGPPGEMGTCF-----KACELHI----- 259
 Db 436 GACAEETGKCECLPRFVGEI-----ASVYCAPKKEEENVNVTITGWLPELQAPKA 444
 Qy 260 -FGWCKPFRSG-----ORCKSYVFLDPPYG-----CSCATGKWKGLQCNHA 301
 Db 496 GFGTGTCTGADGVIYNTAGC-----VECVDAIGSEHNGNSASTGQCKCKPAVAGLSCKR- 551
 Qy 302 CHPGFYGPDKKLECNLN-----NDEMTRPFGVCLSPWAGLQCRP-----ESIPPMTPK 353
 Db 552 GQVYFPGDVKPQNCQWMPFBSVQVQVGTCTCKRSGFADKCLDCLAPFYVYNKAC 610
 Qy 352 IVDLPDHIENSGKTRPKACKSCWPLFTNDEMT 384
 Db 611 ACD-----GAGTISPEVDATSGVCPNINPT 646

RESULT 10
 LMA_DROME STANDARD: PPT: 3712 AA.
 AC Q00174;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Laminin alpha chain precursor.
 GS LANA OR LAMA.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota, Metazoa, Arthropoda, Tracheata, Hexapoda, Insecta.
 OC Pterygota, Neoptera, Endopterygota, Diptera, Brachycera, Muscomorpha,
 OC Phyllophaga, Prosophilidae, Prosophilinae.
 OX NBI_TaxID=7227;
 RN [1]
 RX MEDLINE=93049204, PubMed=1425586,
 RA Kuschel-Gullberg M, Garrison K, Mackrell A J, Fessler L L,
 RA Fessler J H;
 RT "Laminin A chain: expression during Drosophila development and
 RT genomic sequence.";
 FT EMBO J. 11:4519-4527(1992).
 RL [2]
 RN [3]
 RX MEDLINE=94038678, PubMed=8223265;
 RA Henchcliffe C, Garcia-Alonso L, Tang T, Goodman C S;
 RT "Genetic analysis of laminin A reveals diverse functions during
 RT morphogenesis in Drosophila.";
 RL [3]
 RN [3]
 RX MEDLINE=92078147, PubMed=1744083;
 RA Garrison K, Mackrell A J, Fessler J H;
 RT "prosophilin laminin A chain sequence, interspecies comparison, and
 RT domain structure of a major carboxyl portion.";
 PL J. Biol. Chem. 266:22899-22904(1991).
 CC -1- FUNCTION: BINDING TO CELLS VIA A HIGH AFFINITY RECEPTOR. LAMININ
 CC IS THOUGHT TO MEDIATE THE ATTACHMENT, MIGRATION, ORGANIZATION OF
 CC CELLS INTO TISSUES DURING EMBRYONIC DEVELOPMENT BY INTERACTING
 CC WITH OTHER EXTRACELLULAR MATRIX COMPONENTS.
 CC -1- FUNCTION: DIVERSE FUNCTIONS DURING MORPHOGENESIS IN DROSOPHILA.
 CC COMPLETE LOSS-OF-FUNCTION MUTATIONS LEAD TO LATE EMBRYONIC
 CC LETHALITY. CERTAIN PARTIAL LOSS-OF-FUNCTION MUTATIONS GIVE RAISE
 CC TO ESCAPER ADULTS, WHICH HAVE ROUGH EYES ASSOCIATED WITH CHANGES
 CC IN CELL FATE AND PATTERN, MISSHAPEN LENS AND DEFECTS IN WING
 CC STRUCTURE.
 CC -1- SUBUNIT: LAMININ IS A COMPLEX GLYCOPROTEIN, CONSISTING OF THREE
 CC DIFFERENT POLYPEPTIDE CHAINS (ALPHA, BETA, GAMMA), WHICH ARE BOUND
 CC TO EACH OTHER BY DISULFIDE BONDS INTO A CROSS-SHAPED MOLECULE
 CC COMPRISING ONE LONG & THREE SHORT ARMS WITH GLOBULES AT EACH END.
 CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR; FOUND IN THE BASEMENT
 CC MEMBRANES (MAJOR COMPONENT).
 CC -1- TISSUE SPECIFICITY: NEWLY FORMED MESODERM AND LATER PROMINENTLY

CC EXPRESSED IN HEMOCYTES, WHICH ALSO SYNTHESIZE COLLAGEN IV.
 CC -1- DEVELOPMENTAL STAGE: DURING MORPHOGENESIS, MOSTLY IN EMBRYO
 CC DEVELOPMENT, AT 15-16 HOURS.
 CC -1- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT
 CC WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.
 CC -1- DOMAIN: DOMAINS VI, IV AND G ARE GLOBULAR.
 CC -1- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
 CC -1- SIMILARITY: CONTAINS 21.5 LAMININ EGF-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 LAMININ DOMAIN IV (DOMAIN IV) IS NOT
 CC SIMILAR TO LAMININ DOMAIN IV.
 CC -1- SIMILARITY: CONTAINS 5 LAMININ G-LIKE DOMAINS.
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See [http://www.isb.sib.ch/announcements/](http://www.isb.sib.ch/announcements)
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMDL; M96388; AAA28662.1;
 CC EMDL; L07288; AAC37178.1;
 CC EMDL; M75882; AAA28661.1;
 CC HSSP; P02468; LILE.
 CC FlyBase; FBgn0002526; Lama.
 CC InterPro; IPR000581; EGF-like.
 CC InterPro; IPR001886; LamNT.
 CC InterPro; IPR000034; Laminin_B.
 CC InterPro; IPR02049; Laminin_EGF.
 CC InterPro; IPR001791; Laminin_G.
 CC InterPro; IPR001230; Ircy1tn.
 CC Pfam; PF00052; laminin_B; 1.
 CC Pfam; PF00054; laminin_EGF; 20.
 CC Pfam; PF00054; laminin_G; 5.
 CC Pfam; PF00055; laminin_Nterm; 1.
 CC PRINTS; PK00011; EGF/LAMININ.
 CC PRODOM; P002082; LamNT; 1.
 CC PRODOM; P003031; Laminin_B; 1.
 CC SMART; SM00180; EGF_Lam; 17.
 CC SMART; SM00001; EGF-like; 1.
 CC SMART; SM00281; LamB; 1.
 CC SMART; SM00282; LamG; 5.
 CC SMART; SM00136; LamNT; 1.
 CC PROSITE; PS00022; EGF_1; 17.
 CC PROSITE; PS01186; EGF_2; 5.
 CC PROSITE; PS01248; LAMININ_TYPE_EGF; 19.
 CC PROSITE; PS50025; LAM_G_DOMAIN; 5.
 CC Glycoprotein; Basement membrane, Extracellular matrix, Coiled coil;
 CC Laminin EGF-like domain, Cell adhesion, Repeat, Signal.
 CC SIGNAL; 1 22 POTENTIAL.
 CC CHAIN; 23 3712 LAMININ ALPHA CHAIN.
 CC DOMAIN; 25 272 LAMININ N-TERMINAL (DOMAIN VI).
 CC DOMAIN; 273 352 LAMININ EGF-LIKE 1.
 CC DOMAIN; 333 402 LAMININ EGF-LIKE 2.
 CC DOMAIN; 403 447 LAMININ EGF-LIKE 3.
 CC DOMAIN; 448 494 LAMININ EGF-LIKE 4.
 CC DOMAIN; 495 540 LAMININ EGF-LIKE 5.
 CC DOMAIN; 541 586 LAMININ EGF-LIKE 6.
 CC DOMAIN; 587 631 LAMININ EGF-LIKE 7.
 CC DOMAIN; 632 676 LAMININ EGF-LIKE 8.
 CC DOMAIN; 677 731 LAMININ EGF-LIKE 9.
 CC DOMAIN; 732 784 LAMININ EGF-LIKE 10.
 CC DOMAIN; 785 815 LAMININ EGF-LIKE 11 (INCOMPLETE).
 CC DOMAIN; 816 1374 LAMININ EGF-LIKE 12.
 CC DOMAIN; 1375 1420 LAMININ EGF-LIKE 13.
 CC DOMAIN; 1421 1465 LAMININ EGF-LIKE 14.
 CC DOMAIN; 1466 1513 LAMININ EGF-LIKE 15.
 CC DOMAIN; 1514 1564 LAMININ EGF-LIKE 16.
 CC DOMAIN; 1565 1574 LAMININ EGF-LIKE 16 (N-TERMINAL).
 CC DOMAIN; 1575 1775 LAMININ DOMAIN IV (DOMAIN IV).
 CC DOMAIN; 1776 1808 LAMININ EGF-LIKE 16 (C-TERMINAL).
 CC DOMAIN; 1809 1858 LAMININ EGF-LIKE 17.
 CC DOMAIN; 1859 1916 LAMININ EGF-LIKE 18.

1917 1969
 1970 2016
 2017 2063
 2064 2111
 2112 2271
 2272 2368
 2369 2408
 2409 2448
 2449 2487
 2488 2526
 2527 2565
 2566 2604
 2605 2643
 2644 2681
 2682 2720
 2721 2758
 2759 2796
 2797 2834
 2835 2872
 2873 2910
 2911 2948
 2949 2986
 2987 3024
 3025 3062
 3063 3100
 3101 3138
 3139 3176
 3177 3214
 3215 3252
 3253 3290
 3291 3328
 3329 3366
 3367 3404
 3405 3442
 3443 3480
 3481 3518
 3519 3556
 3557 3594
 3595 3632
 3633 3670
 3671 3708
 3709 3746
 3747 3784
 3785 3822
 3823 3860
 3861 3898
 3899 3936
 3937 3974
 3975 4012
 4013 4050
 4051 4088
 4089 4126
 4127 4164
 4165 4202
 4203 4240
 4241 4278
 4279 4316
 4317 4354
 4355 4392
 4393 4430
 4431 4468
 4469 4506
 4507 4544
 4545 4582
 4583 4620
 4621 4658
 4659 4696
 4697 4734
 4735 4772
 4773 4810
 4811 4848
 4849 4886
 4887 4924
 4925 4962
 4963 5000
 5001 5038
 5039 5076
 5077 5114
 5115 5152
 5153 5190
 5191 5228
 5229 5266
 5267 5304
 5305 5342
 5343 5380
 5381 5418
 5419 5456
 5457 5494
 5495 5532
 5533 5570
 5571 5608
 5609 5646
 5647 5684
 5685 5722
 5723 5760
 5761 5798
 5799 5836
 5837 5874
 5875 5912
 5913 5950
 5951 5988
 5989 6026
 6027 6064
 6065 6102
 6103 6140
 6141 6178
 6179 6216
 6217 6254
 6255 6292
 6293 6330
 6331 6368
 6369 6406
 6407 6444
 6445 6482
 6483 6520
 6521 6558
 6559 6596
 6597 6634
 6635 6672
 6673 6710
 6711 6748
 6749 6786
 6787 6824
 6825 6862
 6863 6900
 6901 6938
 6939 6976
 6977 7014
 7015 7052
 7053 7090
 7091 7128
 7129 7166
 7167 7204
 7205 7242
 7243 7280
 7281 7318
 7319 7356
 7357 7394
 7395 7432
 7433 7470
 7471 7508
 7509 7546
 7547 7584
 7585 7622
 7623 7660
 7661 7698
 7699 7736
 7737 7774
 7775 7812
 7813 7850
 7851 7888
 7889 7926
 7927 7964
 7965 8002
 8003 8040
 8041 8078
 8079 8116
 8117 8154
 8155 8192
 8193 8230
 8231 8268
 8269 8306
 8307 8344
 8345 8382
 8383 8420
 8421 8458
 8459 8496
 8497 8534
 8535 8572
 8573 8610
 8611 8648
 8649 8686
 8687 8724
 8725 8762
 8763 8800
 8801 8838
 8839 8876
 8877 8914
 8915 8952
 8953 8990
 8991 9028
 9029 9066
 9067 9104
 9105 9142
 9143 9180
 9181 9218
 9219 9256
 9257 9294
 9295 9332
 9333 9370
 9371 9408
 9409 9446
 9447 9484
 9485 9522
 9523 9560
 9561 9598
 9599 9636
 9637 9674
 9675 9712
 9713 9750
 9751 9788
 9789 9826
 9827 9864
 9865 9902
 9903 9940
 9941 9978
 9979 10016
 10017 10054
 10055 10092
 10093 10130
 10131 10168
 10169 10206
 10207 10244
 10245 10282
 10283 10320
 10321 10358
 10359 10396
 10397 10434
 10435 10472
 10473 10510
 10511 10548
 10549 10586
 10587 10624
 10625 10662
 10663 10700
 10701 10738
 10739 10776
 10777 10814
 10815 10852
 10853 10890
 10891 10928
 10929 10966
 10967 11004
 11005 11042
 11043 11080
 11081 11118
 11119 11156
 11157 11194
 11195 11232
 11233 11270
 11271 11308
 11309 11346
 11347 11384
 11385 11422
 11423 11460
 11461 11498
 11499 11536
 11537 11574
 11575 11612
 11613 11650
 11651 11688
 11689 11726
 11727 11764
 11765 11802
 11803 11840
 11841 11878
 11879 11916
 11917 11954
 11955 11992
 11993 12030
 12031 12068
 12069 12106
 12107 12144
 12145 12182
 12183 12220
 12221 12258
 12259 12296
 12297 12334
 12335 12372
 12373 12410
 12411 12448
 12449 12486
 12487 12524
 12525 12562
 12563 12600
 12601 12638
 12639 12676
 12677 12714
 12715 12752
 12753 12790
 12791 12828
 12829 12866
 12867 12904
 12905 12942
 12943 12980
 12981 13018
 13019 13056
 13057 13094
 13095 13132
 13133 13170
 13171 13208
 13209 13246
 13247 13284
 13285 13322
 13323 13360
 13361 13398
 13399 13436
 13437 13474
 13475 13512
 13513 13550
 13551 13588
 13589 13626
 13627 13664
 13665 13702
 13703 13740
 13741 13778
 13779 13816
 13817 13854
 13855 13892
 13893 13930
 13931 13968
 13969 14006
 14007 14044
 14045 14082
 14083 14120
 14121 14158
 14159 14196
 14197 14234
 14235 14272
 14273 14310
 14311 14348
 14349 14386
 14387 14424
 14425 14462
 14463 14500
 14501 14538
 14539 14576
 14577 14614
 14615 14652
 14653 14690
 14691 14728
 14729 14766
 14767 14804
 14805 14842
 14843 14880
 14881 14918
 14919 14956
 14957 14994
 14995 15032
 15033 15070
 15071 15108
 15109 15146
 15147 15184
 15185 15222
 15223 15260
 15261 15298
 15299 15336
 15337 15374
 15375 15412
 15413 15450
 15451 15488
 15489 15526
 15527 15564
 15565 15602
 15603 15640
 15641 15678
 15679 15716
 15717 15754
 15755 15792
 15793 15830
 15831 15868
 15869 15906
 15907 15944
 15945 15982
 15983 16020
 16021 16058
 16059 16096
 16097 16134
 16135 16172
 16173 16210
 16211 16248
 16249 16286
 16287 16324
 16325 16362
 16363 16400
 16401 16438
 16439 16476
 16477 16514
 16515 16552
 16553 16590
 16591 16628
 16629 16666
 16667 16704
 16705 16742
 16743 16780
 16781 16818
 16819 16856
 16857 16894
 16895 16932
 16933 16970
 16971 17008
 17009 17046
 17047 17084
 17085 17122
 17123 17160
 17161 17198
 17199 17236
 17237 17274
 17275 17312
 17313 17350
 17351 17388
 17389 17426
 17427 17464
 17465 17502
 17503 17540
 17541 17578
 17579 17616
 17617 17654
 17655 17692
 17693 17730
 17731 17768
 17769 17806
 17807 17844
 17845 17882
 17883 17920
 17921 17958
 17959 18000
 18001 18040
 18041 18080
 18081 18120
 18121 18160
 18161 18200
 18201 18240
 18241 18280
 18281 18320
 18321 18360
 18361 18400
 18401 18440
 18441 18480
 18481 18520
 18521 18560
 18561 18600
 18601 18640
 18641 18680
 18681 18720
 18721 18760
 18761 18800
 18801 18840
 18841 18880
 18881 18920
 18921 18960
 18961 19000
 19001 19040
 19041 19080
 19081 19120
 19121 19160
 19161 19200
 19201 19240
 19241 19280
 19281 19320
 19321 19360
 19361 19400
 19401 19440
 19441 19480
 19481 19520
 19521 19560
 19561 19600
 19601 19640
 19641 19680
 19681 19720
 19721 19760
 19761 19800
 19801 19840
 19841 19880
 19881 19920
 19921 19960
 19961 20000
 20001 20040
 20041 20080
 20081 20120
 20121 20160
 20161 20200
 20201 20240
 20241 20280
 20281 20320
 20321 20360
 20361 20400
 20401 20440
 20441 20480
 20481 20520
 20521 20560
 20561 20600
 20601 20640
 20641 20680
 20681 20720
 20721 20760
 20761 20800
 20801 20840
 20841 20880
 20881 20920
 20921 20960
 20961 21000
 21001 21040
 21041 21080
 21081 21120
 21121 21160
 21161 21200
 21201 21240
 21241 21280
 21281 21320
 21321 21360
 21361 21400
 21401 21440
 21441 21480
 21481 21520
 21521 21560
 21561 21600
 21601 21640
 21641 21680
 21681 21720
 21721 21760
 21761 21800
 21801 21840
 21841 21880
 21881 21920
 21921 21960
 21961 22000
 22001 22040
 22041 22080
 22081 22120
 22121 22160
 22161 22200
 22201 22240
 22241 22280
 22281 22320
 22321 22360
 22361 22400
 22401 22440
 22441 22480
 22481 22520
 22521 22560
 22561 22600
 22601 22640
 22641 22680
 22681 22720
 22721 22760
 22761 22800
 22801 22840
 22841 22880
 22881 22920
 22921 22960
 22961 23000
 23001 23040
 23041 23080
 23081 23120
 23121 23160
 23161 23200
 23201 23240
 23241 23280
 23281 23320
 23321 23360
 23361 23400
 23401 23440
 23441 23480
 23481 23520
 23521 23560
 23561 23600
 23601 23640
 23641 23680
 23681 23720
 23721 23760
 23761 23800
 23801 23840
 23841 23880
 23881 23920
 23921 23960
 23961 24000
 24001 24040
 24041 24080
 24081 24120
 24121 24160
 24161 24200
 24201 24240
 24241 24280
 24281 24320
 24321 24360
 24361 24400
 24401 24440
 24441 24480
 24481 24520
 24521 24560
 24561 24600
 24601 24640
 24641 24680
 24681 24720
 24721 24760
 24761 24800
 24801 24840
 24841 24880
 24881 24920
 24921 24960
 24961 25000
 25001 25040
 25041 25080
 25081 25120
 25121 25160
 25161 25200
 25201 25240
 25241 25280
 25281 25320
 25321 25360
 25361 25400
 25401 25440
 25441 25480
 25481 25520
 25521 25560
 25561 25600
 25601 25640
 25641 25680
 25681 25720
 25721 25760
 25761 25800
 25801 25840
 25841 25880
 25881 25920
 25921 25960
 25961 26000
 26001 26040
 26041 26080
 26081 26120
 26121 26160
 26161 26200
 26201 26240
 26241 26280
 26281 26320
 26321 26360
 26361 26400
 26401 26440
 26441 26480
 26481 26520
 26521 26560
 26561 26600
 26601 26640
 26641 26680
 26681 26720
 26721 26760
 26761 26800
 26801 26840
 26841 26880
 26881 26920
 26921 26960
 26961 27000
 27001 27040
 27041 27080
 27081 27120
 27121 27160
 27161 27200
 27201 27240
 27241 27280
 27281 27320
 27321 27360
 27361 27400
 27401 27440
 27441 27480
 27481 27520
 27521 27560
 27561 27600
 27601 27640
 27641 27680
 27681 27720
 27721 27760
 27761 27800
 27801 27840
 27841 27880
 27881 27920
 27921 27960
 27961 28000
 28001 28040
 28041 28080
 28081 28120
 28121 28160
 28161 28200
 28201 28240
 28241 28280
 28281 28320
 28321 28360
 28361 28400
 28401 28440
 28441 28480
 28481 28520
 28521 28560
 28561 28600
 28601 28640
 28641 28680
 28681 28720
 28721 28760
 28761 28800
 28801 28840
 28841 28880
 28881 28920
 28921 28960
 28961 29000
 29001 29040
 29041 29080
 29081 29120
 29121 29160
 29161 29200
 29201 29240
 29241 29280
 29281 29320
 29321 29360
 29361 29400
 29401 29440
 29441 29480
 29481 29520
 29521 29560
 29561 29600
 29601 29640
 29641 29680
 29681 29720
 29721 29760
 29761 29800
 29801 29840
 29841 29880
 29881 29920
 29921 29960
 29961 30000
 30001 30040
 30041 30080
 30081 30120
 30121 30160
 30161 30200
 30201 30240
 30241 30280
 30281 30320
 30321 30360
 30361 30400
 30401 30440
 30441 30480
 30481 30520
 30521 30560
 30561 30600
 30601 30640
 30641 30680
 30681 30720
 30721 30760
 30761 30800
 30801 30840
 30841 30880
 30881 30920
 30921 30960
 30961 31000
 31001 31040
 31041 31080
 31081 31120
 31121 31160
 31161 31200
 31201 31240
 31241 31280
 31281 31320
 31321 31360
 31361 31400
 31401 31440
 31441 31480
 31481 31520
 31521 31560
 31561 31600
 31601 31640
 31641 31680
 31681 31720
 31721 31760
 31761 31800
 31801 31840
 31841 31880
 31881 31920
 31921 31960
 31961 32000
 32001 32040
 32041 32080
 32081 32120
 32121 32160
 32161 32200
 32201 32240
 32241 32280
 32281 32320
 32321 32360
 32361 32400
 32401 32440
 32441 32480
 32481 32520
 32521 32560
 32561 32600
 32601 32640
 32641 32680
 32681 32720
 32721 32760
 32761 32800
 32801 32840
 32841 32880
 32881 3

FT DISULFID 940 941 BY SIMILARITY.
 FT DISULFID 945 950 BY SIMILARITY.
 FT DISULFID 952 961 BY SIMILARITY.
 FT DISULFID 968 979 BY SIMILARITY.
 FT DISULFID 973 988 BY SIMILARITY.
 FT DISULFID 990 999 BY SIMILARITY.
 FT DISULFID 1005 1019 BY SIMILARITY.
 FT DISULFID 1011 1028 BY SIMILARITY.
 FT DISULFID 1040 1049 BY SIMILARITY.
 FT DISULFID 1046 1057 BY SIMILARITY.
 FT DISULFID 1051 1069 BY SIMILARITY.

Query Match: 7.9%; Score 206.5; DB 1; Length 1964;
 Best local Similarity 24.4%; Pred. No. 2,116 88;
 Matches 72; Conservative 37; Mismatches 85; Indels 101; Gaps 19;

QY 211 CEAQKQWQPE-NH-----L-TA-MNN-V-H-ELIG-EYLTPF-FPMKPTAEK----- 253
 DQ 676 CDFGTGTPCEETELGGCISTPCAHGGICHPQPSGNYCTCPAGYMGLTCESEFVTACHSGPC 745
 QY 254 ----A-VEL-----HTFETPK-----EPSSGQGEW-KSYVPV-L--EFGYKES-ATG 272
 DQ 736 LNGGSGSIPPEGYS-TLPESH-GRFQ-TAVH-V-SAS-LNAGT-VNKP-E-IFR-L-ATG 744
 QY 293 WKLCQNEACHGFCYCHUCK-----LRSCNNG-----EMCDR----- 325
 DQ 794 FQGLICEETNSCADSPCKNKATQDTPRGARCLDSEGYIGSSCQLLDLCAKPKCPHT 853
 QY 426 -----FQGLCSQWQGLQ-EPEKTPMTKIVCLPHHFNVS-KENPFI-KASGWP 476
 DQ 854 AKLCQSGPSFO-CLCQWTCALD-----FP-LSCMAAMSQCILHSG-----LCUNGLAC 903
 QY 477 LPTNEMTLVKPPG-----TVHPKFNHIDHPSVAFTTHPTLPDPSGVVV 424
 DQ 904 HTTSSSYFEP-EP-EP-RTQ-NNMP-EPNP-HPS-V-----EPESQ-VVC 950

RESULT 12
 DILL4_MOUSE
 ID DILL4_MOUSE STANDARD: PRT: 686 AA.
 AC Q9J171; Q9J172;
 DT 16-OCT-2001 (rel. 40, Created)
 DE 16-OCT-2001 (rel. 40, last sequence update)
 DE 16-OCT-2001 (rel. 40, last annotation update)
 DE Delta-like protein 4 precursor (Drosophila Delta homolog 4)
 GN DILL4.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Mammalia; Eutheria; Rodentia; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Adipose tissue;
 RX MEDLINE=20296719; PubMed=10847024;
 RA Shutter J.R., Scully S., Fan W., Richards W.G., Kitajewski J.,
 RA Deblandre G.A., Kinter G.W., Stark K.L.;
 RT "DILL4, a novel Notch ligand expressed in arterial endothelium";
 RL Genes Dev. 14:1313-1318(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21064937; PubMed=11144954;
 RA Yoneya T., Ishihara T., Nagao K., Yamada Y., Yamamoto T., Miyatani S.,
 RA Nishikawa M.;
 RT "Molecular cloning of Delta-4, a new mouse and human Notch ligand";
 RL J. Biochem. 129:27-34(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/JOLA;
 RA Mailhos C., Lewis J., Ish-Horowicz D.;
 RT "A novel Notch ligand expressed in embryonic and tumour
 vasculature";
 RL Submitted (MAY-2000) to the EMBL/GenBank/SwissProt databases
 CC FUNCTION: PLAYS A ROLE IN THE NOTCH SIGNALING PATHWAY. ACTIVATES

CC NOTCH-1 AND NOTCH-4.
 CC 1- SURCELLULAR BINDS TO NOTCH-1 AND NOTCH-4.
 CC 1- SUBCELLULAR LOCATION: Type I membrane protein (Probable).
 CC 1- TISSUE SPECIFICITY: EXPRESSED IN VASCULAR ENDOTHELIUM.
 CC 1- SIMILARITY: BELONGS TO THE DELTA/SERRATE/JAGGED FAMILY.
 CC 1- SIMILARITY: BELONGS TO THE DELTA/SERRATE/JAGGED FAMILY.
 CC This SWISS PRO entry is copyrighted. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/annouce/>
 or send an email to license@isb-sib.ch).

EMBL: AF2534459; AAF76428.1;
 EMBL: AF2534454; AAF76785.1;
 EMBL: AF043893; HAH18580.1;
 MGI: 1859388; D114.
 InterPro: IPR001152; Asx_hydroxyl.
 InterPro: IPR001774; DSL.
 InterPro: IPR00561; EGF-like.
 InterPro: IPR00742; EGF_2.
 InterPro: IPR001881; EGF_Ca.
 InterPro: IPR001438; EGF_II.
 InterPro: IPR002049; laminin_PGF.
 Pfam: PF01414; DSL; 1.
 PRINTS: PR00010; EGF000.
 PRINTS: PR00011; EGF000.
 SMART: SM00051; DSL; 1.
 SMART: SM00179; EGF_CA; 3.
 SMART: SM00001; EGF_Like; 5.
 ProSite: PS00010; ASX_HYDROXYL; 1.
 ProSite: PS00022; EGF_1; 8.
 ProSite: PS01186; EGF_2; 7.
 Signal: EGF-like domain; Repeat; Transmembrane; Developmental protein;
 KW Differentiation; Glycoprotein.
 FT SIGNAL 1 26 POTENTIAL.
 FT CHAIN 27 686 DELTA-LIKE PROTEIN 4.
 FT DOMAIN 27 532 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 533 553 POTENTIAL.
 FT DOMAIN 554 686 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 174 218 DELTA-SERRATE-LAC2.
 FT DOMAIN 219 252 EGF-LIKE 1.
 FT DOMAIN 253 283 EGF-LIKE 2.
 FT DOMAIN 285 323 EGF-LIKE 3.
 FT DOMAIN 325 361 EGF-LIKE 4.
 FT DOMAIN 364 401 EGF-LIKE 5.
 FT DOMAIN 403 439 EGF-LIKE 6.
 FT DOMAIN 441 477 EGF-LIKE 7.
 FT DOMAIN 481 519 EGF-LIKE 8.
 FT DISULFID 223 234 BY SIMILARITY.
 FT DISULFID 227 240 BY SIMILARITY.
 FT DISULFID 242 251 BY SIMILARITY.
 FT DISULFID 254 265 BY SIMILARITY.
 FT DISULFID 260 271 BY SIMILARITY.
 FT DISULFID 273 282 BY SIMILARITY.
 FT DISULFID 289 301 BY SIMILARITY.
 FT DISULFID 295 311 BY SIMILARITY.
 FT DISULFID 313 322 BY SIMILARITY.
 FT DISULFID 329 340 BY SIMILARITY.
 FT DISULFID 334 349 BY SIMILARITY.
 FT DISULFID 351 360 BY SIMILARITY.
 FT DISULFID 367 378 BY SIMILARITY.
 FT DISULFID 372 389 BY SIMILARITY.
 FT DISULFID 391 400 BY SIMILARITY.
 FT DISULFID 407 418 BY SIMILARITY.
 FT DISULFID 412 427 BY SIMILARITY.
 FT DISULFID 429 438 BY SIMILARITY.
 FT DISULFID 445 465 BY SIMILARITY.
 FT DISULFID 467 476 BY SIMILARITY.
 FT DISULFID 485 496 BY SIMILARITY.
 FT DISULFID 490 507 BY SIMILARITY.

THE INNER PART OF EMBRYO IS ONE OF THE FIRST STEPS OF CNS DEVELOPMENT IN INSECTS, THIS PROCESS IS UNDER CONTROL OF THE NEUROGENIC GENES.

-!- MISCELLANEOUS: NOTCH AND SERRATE MAY INTERACT AT THE PROTEIN LEVEL, IT IS CONCEIVABLE THAT THE SERRATE AND DELTA PROTEINS MAY COMPETE FOR BINDING WITH THE NOTCH PROTEIN.

-!- SIMILARITY: CONTAINS 14 EGF-LIKE DOMAINS.

-!- SIMILARITY: BELONGS TO THE DELTA/SERRATE/JAGGED FAMILY.

This SWISS PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/>) or send an email to license@isb.sib.ch.

CC	EMBL; X56811; CAA40148.1; ?;
CC	EMBL; M35759; AA28938.1; ?;
DR	PIR; A46666; A46666;
DR	PIR; S16878; S16878;
DR	HSSP; P00743; ICGF.
DR	FlyBase; Fg00004197; Ser.
DR	InterPro; IPRO00152; Asx_Hydroxyl.
DR	InterPro; IPRO01774; DSL.
DR	InterPro; IPRO00561; EGF-like.
DR	InterPro; IPRO00742; EGF_2.
DR	InterPro; IPRO01881; WGF_Ca.
DR	InterPro; IPRO01438; EGF_11.
DR	Flam; PF01414; DSL; 1.
DR	Flam; PF00008; EGF; 11.
DR	PRINTS; PR00010; EGFBLOOD.
DR	SMART; SM00051; DSL; 1.
DR	SMART; SM00179; RGF_CA; 7.
DR	SMART; SM00001; RGF_Like; 5.
DR	PROSITE; PS00010; ASX_HYDROXYL; 7.
DE	PROSITE; PS00322; EGF_1; 14.
DE	PROSITE; PS01186; EGF_2; 8.
KW	Differentiation, Repeat, RGF-like domain, Transmembrane;
KW	Glycoprotein; Signal.
FT	SIGNAL 1 83
FT	CHAIN 84 1408
FT	DOMAIN 84 1223
FT	TRANSMEM 1224 1249
FT	DOMAIN 1250 1408
FT	DOMAIN 284 317
FT	DOMAIN 315 349
FT	DOMAIN 351 389
FT	DOMAIN 391 489
FT	DOMAIN 407 476
FT	DOMAIN 491 527
FT	DOMAIN 529 609
FT	DOMAIN 611 646
FT	DOMAIN 648 684
FT	DOMAIN 686 721
FT	DOMAIN 723 797
FT	DOMAIN 737 769
FT	DOMAIN 799 835
FT	DOMAIN 837 877
FT	DOMAIN 879 915
FT	DOMAIN 917 953
FT	DISULFID 288 299
FT	DISULFID 292 305
FT	DISULFID 307 316
FT	DISULFID 319 330
FT	DISULFID 325 337
FT	DISULFID 339 348
FT	DISULFID 355 367
FT	DISULFID 361 377
FT	DISULFID 379 388
FT	DISULFID 395 406
FT	DISULFID 400 477

Query Match
Best Local Similarity 7.7%; Score 199.5; DB 1; Length 686;
Matches 50; Conservative 24.8%; Pred. No. 2.2e+08;
Mismatches 17; Mismatches 48; Idents 87; Gaps 9;

QY	211	*EAQKWEFEENHILAT-----ACMNKV-----	GHEDG... 238
DB	176	CSDNYVRESPLCKRKRRHFHVVECPENSLNLPDWIKKYLDQPCLSAHEQNGYS	235
QY	239	---ECIGPGHWGRICPKACAFHIFEGRICKEKSCQEDSKSYVPELDPDYQSCAYWKGC	295
DB	236	KPEFLCRPGWGVRPNF-----ITHFGHPRHTST--PWLA'DEWSG	279
QY	296	LQCNE-----ACHFYGYDECKL---P-S---NNNGEM	322
DB	280	LFCDQGNVCITHRSCAKNGSICSNSGRKGVCICGLPGVICHHCLHGLSKASNIKNNGGS	339
QY	323	---DFQGCLCSPGWQLQCE 341	
DB	340	KXDQNSHYCLCPGGYGQHCE 361	

RESULT 13

ID	SEPP_DROME	STANDARDP:	PRT: 1408 AA
OC	P81668:		
DI	01-NOV-1990 (Rel. 16, Created)		
DT	01-JUL-1993 (Rel. 25, last sequence update)		
DT	15-JUL-1999 (Rel. 38, last annotation update)		
DE	Serrate protein precursor (Beaded protein).		
GN	SER OR BD.		
OS	Drosophila melanogaster (fruit fly)		
OC	Eukaryota; Metazoa; Arthropoda; Insecta; Hexapoda; Insecta;		
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
OC	Ephydroidea; Drosophilidae; Drosophilla.		
OX	NCBI_Taxid=7227;		
RI	[1]		
RC	SEQUENCE FROM N.A.		
RC	STRAIN=ORECON-R;		
RX	MEDLINE=91347903; PubMed=1840519;		
RA	Thomas U., Speicher S.A., Knust K.;		
RT	"The Drosophila gene Serrate encodes an EGF-like transmembrane		
RT	protein with a complex expression pattern in embryos and wing		
RT	discs.";		
RL	Development 111:749-761(1991).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=91094666; PubMed=2125287;		
RA	Fleming R.J., Scottale T.N., Diederich R.J., Artavanis-Tsakonas S.;		
RT	"The gene Serrate encodes a putative EGF-like transmembrane protein		
RT	essential for proper ectodermal development in drosophila		
RT	melanogaster.";		
RL	Genes Dev. 4:2188-2201(1990).		

-!- FUNCTION: ESSENTIAL FOR PROPER ECTODERMAL DEVELOPMENT. SERRATE MAY REPRESENT AN ELEMENT IN A NETWORK OF INTERACTING MOLECULES OPERATING AT THE CELL SURFACE DURING THE DIFFERENTIATION OF CERTAIN TISSUES.

-!- SUBCELLULAR LOCATION: Type I membrane protein.

-!- TISSUE SPECIFICITY: APPEARS TO BE RESTRICTED EXCLUSIVELY TO CELLS OF ECTODERMAL ORIGIN.

-!- MISCELLANEOUS: SEPARATION OF NEURON GLASTS FROM THE ECTODERM INTO

Db	1067	-----RRVAPPELPAKSSCW 1081
RESULT	14	
WIFI_HUMAN		
ID	WIFI_HUMAN	STANDARD: PRT: 379 AA.
AC	Q9Y5W5	
DT	16-OCT-2001	(Rel. 40, Created)
DI	16-OCT-2001	(Rel. 40, last sequence update)
DI	16-OCT-2001	(Rel. 40, last annotation update)
DE	Wnt inhibitory factor 1 precursor (WIF-1).	
GN	WIF1	
OS	Homo sapiens (Human).	
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;	
CC	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.	
OX	NCBI_TaxID=9606;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
EX	MEDLINE=99215557; PubMed=10201374;	
RA	Hsieh J.-C., Kodjabachian L., Reibert M.L., Ratiner A.,	
RA	Smallwood P.M., Samos C.H., Nusse R., David J.H., Nathans J.;	
RT	"A new secreted protein that binds to Wnt proteins and inhibits their	
RT	activities."	
RL	Nature 398:431-436(1999).	
CC	-!- FUNCTION: BINDS TO WNT PROTEINS AND INHIBITS THEIR ACTIVITIES. MAY	
CC	BE INVOLVED IN MESODERM SEGMENTATION.	
CC	-!- SUBCELLULAR LOCATION: Secreted.	
CC	-!- SIMILARITY: CONTAINS 5 EGF-LIKE DOMAINS.	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation	
CC	at the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (See http://www.isb-sib.ch/announcements	
CC	or send an email to license@isb-sib.ch).	
EMBL	AF122922; AAC15492.1;	
HSSP	P00743; 10CF.	
DR	MM: 605186;	
DR	InterPro: IPR000561; EGF-like.	
DR	InterPro: IPR002049; Laminin_EGF.	
DR	InterPro: IPR003306; WIF.	
DR	Pfam: PF00008; EGF; 5.	
DR	Pfam: PFG019; WIF; 1.	
DR	PRINTS: PR00011; EGF-LAMININ.	
DR	SMART: SM00181; EGF; 5.	
DR	SMART: SM00469; WIF; 1.	
DR	PROSITE: PS00022; EGF_1; 5.	
DR	PROSITE: PS01186; EGF_2; 4.	
KW	Repeat, EGF like domain, Signal, Developmental protein.	
FT	SIGNAL	1 28
FT	POTENTIAL.	
FT	CHAIN	29 379
FT	WNT INHIBITORY FACTOR 1.	
FT	DOMAIN	177 208
FT	EGF-LIKE 1.	
FT	DOMAIN	209 240
FT	EGF-LIKE 2.	
FT	DOMAIN	241 272
FT	EGF-LIKE 3.	
FT	DOMAIN	273 304
FT	EGF-LIKE 4.	
FT	DOMAIN	305 336
FT	EGF-LIKE 5.	
FT	DISULFID	177 186
FT	POTENTIAL.	
FT	DISULFID	182 192
FT	POTENTIAL.	
FT	DISULFID	198 200
FT	POTENTIAL.	
FT	DISULFID	209 218
FT	POTENTIAL.	
FT	DISULFID	214 224
FT	POTENTIAL.	
FT	DISULFID	233 252
FT	POTENTIAL.	
FT	DISULFID	241 250
FT	POTENTIAL.	
FT	DISULFID	246 256
FT	POTENTIAL.	
FT	DISULFID	262 264
FT	POTENTIAL.	
FT	DISULFID	273 282
FT	POTENTIAL.	
FT	DISULFID	278 288
FT	POTENTIAL.	
FT	DISULFID	294 296
FT	POTENTIAL.	
FT	DISULFID	305 314
FT	POTENTIAL.	
FT	DISULFID	310 320
FT	POTENTIAL.	
FT	DISULFID	326 328
FT	POTENTIAL.	

FT CAPB01VD 88 N-LINKED (GLNAC...) (POTENTIAL)
 FT CAPB01VD 245 N-LINKED (GLNAC...) (POTENTIAL)
 SQ SEQUENCE 379 AA: 41512 MW: 27782376264784 C6C64.

Query Match 7.6%; Score 198.5; DB 1; Length 379;
 Best Local Similarity 26.4%; Ref No 1; Gap 15;
 Matches 72; Conservative 23; Mismatches 95; Indels 81; Gaps 15;

QY 92 KASKINGAYCFGRVRC-PAIRIPIMKMQQASFLPATITMTIVDKNDVNISFKKVLKE 150
 Db 130 KASVQVQGFPLCKQGVAAFEVDVIMNSEGN-----TILQTPQ-----NALIFKICLOA 180
 QY 151 EDVIYKNGSEIISVPHIEVDPILEVILPHAPQDQAGVSYIGNLTSAFTLIVRR 210
 Db 181 PCQGGGNGGFCNFRICECPD--GFHGH----- 208
 QY 211 CEAGKWKPECHILACT-ACMNGVCHEDTGEICGPRGEMGTCEKA-CELIHF-GRTC--K 265
 Db 269 CE-----KALCTPP-AMNGLP-VLPGLP-CPGDFYVNTKAM-STI-FNGGTCTYP 259
 QY 266 FRC-----SDQPRCKSYVCPPIIP-----YQPSCATGKGIQNP-ACHQPGVQDPC 311
 Db 260 GKCTGPELELEQETISKPPFNAPNKKGLGSKCKKSGYQDGLCSKPVCEFG----- 313
 QY 312 KLPSCNNGEMCEPFGQCLSPWQGLQTER 342
 Db 314 -----COAGHTCHPKNCKQCGQGHGHCNK 339

RESULT 15
 DL_DROME
 ID DL_DROME: STANDARD: PRT: 833 AA.
 AC P10041: Q9VUY2: Q9VUY2: 833 AA.
 DT 01-MAR-1989 (Rel. 10, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Neurogenic locus Delta protein precursor.
 GN DL OR C33619.
 OS Drosophila melanogaster (Fruit fly).
 OC Insecta, Hexapoda, Insecta.
 OC Ephydroidea, Ephydroidea, Diptera; Brachyoptera; Muscomorpha;
 OC NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A., TISSUE SPECIFICITY, AND DEVELOPMENTAL STAGE.
 RC TISSUE=Embryo;
 RA Vaessin H., Bremer K.A., Knust E., Campos-Ortega J.A.;
 RT "The neurogenic gene Delta of Drosophila melanogaster is expressed in
 neurogenic territories and encodes a putative transmembrane protein
 with EGF-like repeats.";
 RI EMBO J. 6:3431-3440(1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPAIN-OREGAN P. TISSUE=Embryo,
 RX MEDLINE=89196890; PubMed=149249;
 RA Kopynski C., Alton A., Reitel K., Koch P.J., Maskavitch M.A.T.,
 RT "Delta, a Drosophila neurogenic gene, is transcriptionally complex and
 encodes a protein related to blood coagulation factors and epidermal
 growth factor of vertebrates.";
 RI Genes Dev. 2:1723-1735(1988).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RX MEDLINE=20196036; PubMed=10731132;
 RA Adams M.D., Gelink S.E., Hill P.A., Evans C.A., Gwynne T.D.,
 RA Ananides P.G., Scherer S.E., Li P.W., Hoskins P.A., Galle P.F.,
 RA George B.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortman J.P., Vandell M.D., Zhang Q., Chen L.X.,
 RA Brandon P.C., Rogers Y.H.C., Blarney P.C., Chang M., Pfeiffer H.P.,
 RA Wan K.H., Doyle C., Huxley P.C., Heit G., Nelson C.K., Miklos G.L.G.,
 RA April J.F., Agbayani A., An H.-J., Andrews-Plamkovich R., Baldwin B.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

Reeson K.Y., Benes P.V., Beiman B.P., Bhandari D., Boislakov S.,
 Borkova D., Botchan M.P., Bouck J., Brokstein P., Broitler P.,
 Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 Cherry J.M., Cawley S., Dahlke C., Davenport L.H., Davies P.,
 de Paolis R., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 Dodson K., Doup L.E., Downes M., Duquenois S., Dunkov B.C., Dunn P.,
 Durbin K.J., Evangelista C.C., Ferrar C., Ferris S., Fleischmann W.,
 Fuster G., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 Hostin D., Houston K.A., Howland T.J., Wei M.H., Ikenwale C.,
 Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
 Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 Liu X., Mattei H., McIntosh T.C., McLeod M.P., McPherson D.,
 Mckelvey G., Milshina N.V., Mobarry C., Morris J., Mushrefi A.,
 Mount S.M., Moy M., Murphy R., Murphy L., Muzny D.M., Nelson D.L.,
 Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,
 Palazzolo M., Pittman G.S., Fan S., Pollard J., Puri V., Reese M.G.,
 Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 Spier P., Spradling A.C., Stapleton M., Strong R., Sun E.,
 Svirska K., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 Ye J., Yeh P.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 Zheng X.H., Zhou F.N., Zhou W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RI Science 287:2185-2195(2009).
 PI [4]
 RP SEQUENCE OF 422 621 FROM N.A.
 RC TISSUE=Embryo;
 RX MEDLINE=87218537; PubMed=3107986;
 RA Knust E., Dietrich U., Tepass U., Bremer K.A., Weigel D., Vaessin H.,
 Campos-Ortega J.A.;
 RT "ECF homologous sequences encoded in the genome of Drosophila
 melanogaster, and their relation to neurogenic genes.";
 RI EMBO J. 6:761-766(1987).
 RN [5]
 RP PATTERN OF TRANSCRIPTION, AND CHARACTERIZATION.
 RX MEDLINE=91299246; PubMed=2128477;
 RA Baenlin M., Kratamatschek B., Campos-Ortega J.A.;
 RT "The pattern of transcription of the neurogenic gene Delta of
 Drosophila melanogaster.";
 RI Development 110:905-914(1990).
 CC [1-] FUNCTION: ESSENTIAL FOR PROPER DIFFERENTIATION OF ECTODERM. DL
 CC IS REQUIRED FOR THE CORRECT SEPARATION OF NEURAL AND EPIDERMAL
 CC CELL LINEAGES.
 CC [1-] SUBCELLULAR LOCATION: Type I membrane protein.
 CC ABILITIES, FOR EXAMPLE THE NEUROGENIC ECTODERM AND THE PRIMITIVE
 CC OF THE SENSE ORGANS. LATER EXPRESSION IS RESTRICTED TO THOSE CELLS
 CC THAT HAVE ADOPTED A NEURAL FATE.
 CC [1-] DEVELOPMENTAL STAGE: EXPRESSED BOTH MATERNALLY AND ZYGOTICALLY.
 CC EXPRESSION IS HIGHEST EARLY IN EMBRYONIC DEVELOPMENT (STAGE 5) AND
 CC REDUCES TO A LOW LEVEL DURING LARVAL STAGES.
 CC [1-] MISCELLANEOUS: SEPARATION OF NEUROBLASTS FROM THE ECTODERM INTO
 CC THE INNER PART OF EMBRYO IS ONE OF THE FIRST STEPS OF CNS
 CC DEVELOPMENT IN INSECTS, THIS PROCESS IS UNDER CONTROL OF THE
 CC NEUROGENIC GENES.
 CC [1-] MISCELLANEOUS: NOTCH AND SERRATE MAY INTERACT AT THE PROTEIN
 CC LEVEL, IT IS CONCEIVABLE THAT THE SERRATE AND DELTA PROTEINS MAY
 CC COMPLETE FOR BINDING WITH THE NOTCH PROTEIN.
 CC [1-] SIMILARITY: CONTAINS 9 EGF-LIKE DOMAINS.
 CC [1-] SIMILARITY: BELONGS TO THE DELTA/SERRATE/LAEGED FAMILY.
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.slb.ch/announce/>)

GenCore version 4.5

Copyright (c) 1993 - 2000 CompuGen Ltd

OM protein - protein search, using sw model

Run on: September 23, 2002, 22:14:44, Search time: 31.84 seconds
(without alignments)
2564,498 Million cell updates

Title: US-09-733-764-2 COPY 1 472

Perfect score 26.05

Sequence: 1 MDSLASLVCCVSHLSGTV

VITICULTURAL INSECTARY 172

Scoring table:

Gapop 10.0 , Gapext 0.5

0
1
2
3
4
5
6
7
8
9
A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z

Searched _____ Indexed _____
Serials Acq. 10004920

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Minimum DB seq length:	0
Maximum DB seq length:	200000000

Post-processing: Minimum Match 0.8

Maximum Match 100%
Listing first 45 summaries

Database :

- ```

SP_Knechtig_197:
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phase:*
10: sp.plant:*
11: sp.sporent:*
12: sp.virus:*
13: sp.vvertebrate:*
14: sp.unclassified:*
15: sp.virus:*
16: sp.vvertebrate:*
17: sp.archaea:*

```

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

| Result No. | Query |       | Length | DB | ID     | Description        |
|------------|-------|-------|--------|----|--------|--------------------|
|            | Score | Match |        |    |        |                    |
| 1          | 225R  | 86 7  | 1083   | 13 | Q9WZ24 | Q9WZ24 rattus sp.  |
| 2          | 89G   | 34 5  | 1116   | 13 | Q73791 | Q73791 brachydanio |
| 3          | 327.5 | 12.6  | 434    | 5  | Q9WDA0 | Q9WDA0 drosophila  |
| 4          | 321   | 12.3  | 1664   | 5  | Q91VQ2 | Q91VQ2 caenorhabdi |
| 5          | 317.5 | 12.2  | 1574   | 11 | Q88Z81 | Q88Z81 rattus norv |
| 6          | 316   | 12.1  | 1140   | 4  | Q96K67 | Q96K67 homo sapien |
| 7          | 314   | 12.1  | 969    | 4  | Q96K96 | Q96K96 homo sapien |
| 8          | 289   | 11.1  | 830    | 4  | Q14162 | Q14162 homo sapien |
| 9          | 289   | 11.1  | 830    | 4  | Q43701 | Q43701 homo sapien |
| 10         | 287.5 | 11.0  | 594    | 5  | Q9Y151 | Q9Y151 drosophila  |
| 11         | 286.5 | 11.0  | 594    | 5  | Q9WDA1 | Q9WDA1 drosophila  |
| 12         | 284.5 | 10.9  | 220    | 11 | Q63404 | Q63404 rattus norv |
| 13         | 280   | 10.7  | 1111   | 5  | Q9XWD6 | Q9XWD6 caenorhabdi |
| 14         | 253.5 | 9.7   | 153    | 4  | Q75095 | Q75095 homo sapien |
| 15         | 221   | 8.5   | 117    | 6  | Q9BDX0 | Q9BDX0 sus scrofa  |
| 16         | 221   | 8.5   | 3704   | 5  | P91904 | P91904 caenorhabdi |

|    |       |     |      |    |        |
|----|-------|-----|------|----|--------|
| 17 | 215   | 8.3 | 1512 | 13 | Q9DF36 |
| 18 | 214.5 | 8.2 | 2447 | 13 | Q13149 |
| 19 | 213   | 8.2 | 1476 | 13 | Q9285  |
| 20 | 210   | 8.1 | 598  | 4  | Q6636  |
| 21 | 207   | 7.9 | 3712 | 5  | Q9VRW0 |
| 22 | 206   | 7.9 | 1213 | 13 | Q9DY54 |
| 23 | 205.5 | 7.9 | 1203 | 11 | Q06008 |
| 24 | 205.5 | 7.9 | 1364 | 11 | Q35442 |
| 25 | 205   | 7.9 | 2470 | 11 | Q35516 |
| 26 | 204.5 | 7.9 | 1214 | 13 | Q9DY52 |
| 27 | 204.5 | 7.9 | 2346 | 11 | Q9JL01 |
| 28 | 204   | 7.8 | 762  | 13 | Q43373 |
| 29 | 204.5 | 7.8 | 1370 | 5  | Q9GPN0 |
| 30 | 204.5 | 7.8 | 2471 | 4  | Q04721 |
| 31 | 202.5 | 7.8 | 1308 | 5  | Q9GPM8 |
| 32 | 202.5 | 7.8 | 2471 | 4  | Q9H240 |
| 33 | 200.5 | 7.7 | 1193 | 13 | Q90819 |
| 34 | 200.5 | 7.7 | 1515 | 13 | Q9DF37 |
| 35 | 200.5 | 7.7 | 2471 | 11 | Q90830 |
| 36 | 200.5 | 7.7 | 2824 | 13 | Q9W7R3 |
| 37 | 199.5 | 7.7 | 686  | 11 | Q9DB09 |
| 38 | 199.5 | 7.7 | 1254 | 13 | Q9YH92 |
| 39 | 199.5 | 7.7 | 1254 | 13 | Q9DY56 |
| 40 | 199.5 | 7.7 | 1404 | 5  | Q9VB65 |
| 41 | 199   | 7.6 | 721  | 5  | Q9SVZ0 |
| 42 | 199   | 7.6 | 1212 | 13 | Q42347 |
| 43 | 199   | 7.6 | 2715 | 11 | Q9WTS6 |
| 44 | 198.5 | 7.6 | 464  | 5  | Q95PM9 |
| 45 | 198   | 7.6 | 1242 | 13 | Q9DY57 |

## ALIGNMENT

|                   |                                                       |  |
|-------------------|-------------------------------------------------------|--|
| RESULT            | 1                                                     |  |
| Q90W24            | PRELIMINARY; PR1: 1083 AA.                            |  |
| Q90W24            |                                                       |  |
| AC                | Q90W24;                                               |  |
| DT                | 01-MAY-2000 (TREMBLrel. 13, Created)                  |  |
| DT                | 01-MAY-2000 (TREMBLrel. 13, Last sequence update)     |  |
| DT                | 01-DEC-2000 (TREMBLrel. 19, Last annotation update)   |  |
| DE                | TIE-2-RECEPTOR-LIKE TYROSINE KINASE.                  |  |
| OS                | Rattus sp.                                            |  |
| OC                | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata;   |  |
| OC                | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae   |  |
| NCBI_TaxID=10118; |                                                       |  |
| RN                | [1]                                                   |  |
| RP                | SEQUENCE FROM N.A.                                    |  |
| RP                | MDLINL-94275658; PubMed-7084840;                      |  |
| FA                | "Distinctive p. 55. Goldfarb M., Vankegenure G.D., G. |  |
| FA                | "Distinct rat genes with related profiles of expres   |  |
| RT                | receptor tyrosine kinase family.";                    |  |
| RT                | Oncogene 8;1631-1637(1993).                           |  |
| DR                | HSP; p11362; IFGK.                                    |  |
| DR                | InterPro: IPR000561; EGF-like.                        |  |
| DR                | InterPro: IPR000719; Euk_pkinase.                     |  |
| DR                | InterPro: IPR003961; FN.III.                          |  |
| DR                | InterPro: IPR003549; Laminin_EGF.                     |  |
| DR                | InterPro: IPR001245; Tyr_pkinase.                     |  |
| DR                | Pfam: PF00041; fn3; 3.                                |  |
| DR                | Pfam: PF00069; pkinase; 1.                            |  |
| DR                | PRINTS: PR00011; EGFAMIN.                             |  |
| DR                | PRINTS: PF00109; TYPKINASE.                           |  |
| DR                | SMART: SM00180; EGF_Lam; 1.                           |  |
| DR                | SMART: SM00041; EGF_Like; 1.                          |  |
| DR                | SMART: SM00040; FN3; 3.                               |  |
| DR                | SMART: SM00219; TyPK; 1.                              |  |
| DR                | SMART: PS00622; EGF_1; UNKNOWN_3.                     |  |
| DR                | PROSITE: PS01146; EGF_2; 3.                           |  |
| DR                | PROSITE: PS00137; PROTEIN_KINASE_ATP; 1.              |  |
| DR                | PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.              |  |
| DR                | PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.              |  |
| DR                | ATP binding, EGF-like domain, Glycoprotein, Transf    |  |

KW Tyrosine-protein kinase.  
 SQ SEQUENCE 1083 AA: 12202P MW: 649750 Da PIR4430 CR644.

Query Match 86.7%; Score 2258; DB 11; Length 1083.  
 Best Local Similarity 93.1%; Pred. No. 2.7e-211;  
 Matches 403; Conservative 12; Mismatches 18; Indels 0; Gaps 0;

QY 40 TSLTCLASWFFHEPTITSPDEALMNQJGPEVLTGVTEPWAKKVVVKKKASKINGA 99  
 DB 1 TSLTCLASWFFHEPTITSPDEALMNQJGPEVLTGVTEPWAKKVVVKKKASKINGA 60

QY 100 YFCGVKRGFATIRIRMMKQKQASHLPATLIMIVUKGHNVNISKKVLJKEEDAVLYKNG 159  
 DB 61 YFCGVKRGFATIRIRMMKQKQASHLPATLIMIVUKGHNVNISKKVLJKEEDAVLYKNG 120

QY 160 SFTHSVPRHEVDLLEVLPHAPQPDQAVYSARYVGGNLFISAPTRILVPRCEAKKWPPE 219  
 DB 121 SFTHSVPRHEVDLLEVLPHAPQPDQAVYSARYVGGNLFISAPTRILVPRCEAKKWPPE 180

QY 220 CNHLCTAAMNNVCHFDLSEVTPPEFMSPTEFANTELHTEPTKPEKSGSEPTKSVVF 279  
 DB 181 CNHPCTTKNNVCHFDLSEVTPPEFMSPTEFANTELHTEPTKPEKSGSEPTKSVVF 240

QY 280 CLDPDPGSCATGKGLQNFANRHPGSPGSPGKQKSTNNPMMTQTPGTLSPGSGGLQ 339  
 DB 241 CLDPDPGSCATGKGLQNFANRHPGSPGSPGKQKSTNNPMMTQTPGTLSPGSGGLQ 300

QY 340 CEGEGIPRMTPTKIVDLPDHIENYSKFNPTCKASGWLPTNEEMTLVKPQGVTLBPKNFN 399  
 DB 301 CEKEGIPRMTPTKIVDLPDHIENYSKFNPTCKASGWLPTNEEMTLVKPQGVTLBPKNFN 360

QY 400 HTDFSVVAIFTHIRLPDPSGVWCVSNTVAGVKEKPNISVKVLPKPLNAPNYDITGHN 459  
 DB 361 HTDFSVVAIFTHIRLPDPSGVWCVSNTVAGVKEKPNISVKVLPKPLNAPNYDITGHN 420

QY 460 FAVINISSEPEYFG 472  
 DB 421 FAVINISSEPEYFG 433

RESULT 2  
 ID 073791 PRELIMINARY; PRI: 1116 AA.

AC 073791;  
 DT 01-AUG-1998 (Tremblrel. 07, Created)  
 DT 01-AUG-1998 (Tremblrel. 07, Last sequence update)  
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
 DE ANGIOPOIETIN 1 RECEPTOR PRECURSOR (PC 2.7.1.112) (TYROSINE-PROTEIN KINASE RECEPTOR TIE-2) (ENDOTHELIAL-SPECIFIC RECEPTOR TYROSINE KINASE TIE-2)  
 GN TIE2 (P TIE2 (P TIE-2)  
 OS Brachydanio rerio (zebrafish) (Zebra danio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.  
 OX NCBI\_TaxID:7955;  
 RN [1]  
 RP SEQUENCE FROM N.A. AND TISSUE SPECIFICITY  
 RC TISSUE=EMBRYO;  
 RA MEDLINE-98264593; PubMed-9603430;  
 RA Lyons M.S., Bell R., Stainer D., Peters K.G.;  
 RT "Isolation of the zebrafish homologues for the tie-1 and tie-2 endothelium-specific receptor tyrosine kinases.";  
 RL Dev. Dyn. 212:133-140(1998).  
 CC -!- FUNCTION: THIS PROTEIN IS A PROTEIN TYROSINE KINASE. TRANSMEMBRANE RECEPTOR FOR ANGIOPOIETIN 1. PROBABLY REGULATES ENDOTHELIAL CELL PROLIFERATION, DIFFERENTIATION AND GUIDES THE PROPER PATTERNING OF ENDOTHELIAL CELLS DURING BLOOD VESSEL FORMATION (BY SIMILARITY).  
 CC -!- CATALYTIC ACTIVITY: ATP + A PROTEIN TYROSINE - ADP + PROTEIN TYROSINE PHOSPHATE.  
 CC -!- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.  
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN MOST POPULATIONS OF ENDOTHELIAL

CC 1 CELLS IN 24 HOUR EMBRYOS. NOT PRESENT IN INTERSEGMENTAL VESSELS.  
 CC 1 SIMILARITY TO OTHER PROTEIN TYROSINE KINASES IN THE CATALYTIC DOMAIN.  
 CC -!- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.  
 CC -!- SIMILARITY: CONTAINS 3 FIBRONECTIN TYPE III-LIKE DOMAINS.  
 DR EMBL: AF053632; AAF04331 1; 2;  
 DR HSSD: P11362; 1PKC;  
 DR 2FIN: Z59-GENE-990415-56; tie2;  
 DR InterPro: IPR000561; EGF-like;  
 DR InterPro: IPR000719; EGF\_kinase;  
 DR InterPro: IPR003961; FN\_III;  
 DR InterPro: IPR003599; IG;  
 DR InterPro: IPR001245; Tyr\_kinase;  
 DR Pfam: PF00041; fn3; 3;  
 DR Pfam: PF00069; kinase; 1;  
 DR PRINTS: PR00109; TYRKINASE;  
 DR SMART: SM00181; EGF; 2;  
 DR SMART: SM00050; FN3; 2;  
 DR SMART: SM00409; IG; 1;  
 DR SMART: SMO219; TyreK; 1;  
 DR PROSITE: PS00022; EGF\_1; DNKNWL\_2;  
 DR PROSITE: PS01186; EGF\_2;  
 DR PROSITE: PS00107; PROTEIN\_KINASE\_ATP; 1;  
 DR PROSITE: PS00011; PROTEIN\_KINASE\_DOM; 1;  
 DR PROSITE: PS00109; PROTEIN\_KINASE\_TYR; 1;  
 KW Receptor; Tyrosine-protein kinase; Transmembrane; Signal; ATP-binding; Repeat; EGF-like domain; Transmembrane; Immunoglobulin domain; Glycoprotein; Phosphorylation; Multigene family.  
 KW SIGNAL: 1 21 POTENTIAL;  
 FT CHAIN: 1 21 ANGIOPOIETIN 1 RECEPTOR.  
 FT DOMAIN: 22 1116 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 22 745 POTENTIAL.  
 FT DOMAIN: 746 1116 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN: 767 1116 IG-LIKE C2-TYPE DOMAIN.  
 FT DOMAIN: 45 105 EGF-LIKE 1;  
 FT DOMAIN: 214 256 EGF-LIKE 2;  
 FT DOMAIN: 258 302 EGF-LIKE 3;  
 FT DOMAIN: 304 342 IG-LIKE C2-TYPE DOMAIN.  
 FT DOMAIN: 368 422 FIBRONECTIN TYPE-III.  
 FT DOMAIN: 442 535 FIBRONECTIN TYPE-III.  
 FT DOMAIN: 538 627 FIBRONECTIN TYPE-III.  
 FT DOMAIN: 631 727 PROTEIN KINASE.  
 FT DOMAIN: 816 1088 ATP (BY SIMILARITY).  
 FT NP\_BIND: 822 830 ATP (BY SIMILARITY).  
 FT BINDING: 847 847 BY SIMILARITY.  
 FT ACT\_SITE: 956 956 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).  
 FT MOD\_RES: 984 984 PHOSPHORYLATION (AUTO-) (POTENTIAL).  
 FI MOD\_RES: 1094 1094 PHOSPHORYLATION (AUTO-) (POTENTIAL).  
 FT MOD\_RES: 1105 1105 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD: 110 110 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD: 143 143 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD: 223 223 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD: 367 367 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD: 387 387 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD: 425 425 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD: 590 590 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD: 637 637 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT CARBOHYD: 642 642 N-LINKED (GLCNAC...) (POTENTIAL).  
 SQ SEQUENCE 1116 AA: 122360 MW: 441480 Da PIR4430 CR644.

Query Match 34.5%; Score 899.5; DB 13; Length 1116;  
 Best Local Similarity 38.3%; Pred. No. 1.3e-78;  
 Matches 188; Conservative 88; Mismatches 172; Indels 43; Gaps 15;

QY 1 MSLASIVAGVSLISGTVGAMBLIILINSILPVSQ-AETSLRCLASGW----- 49  
 DB 4 LUSCHALLIG--CWSGSAVKLSIVHIVNIPVVSPLTSLICVSSWSGSGSVLAIG 61

QY 50 ----RPH-EPITIGDFEALMNQDPLEVTDVTRWAKKVVVKKKASKINAYFCG 104  
 DB 62 QEFPRPQGSVIALGQFF-----PHTEPHPHAAATVWSSR-----SHAFGAPYCOI 108



Submitted (MAY 1999) to the EMBL/GenBank/DBJ databases  
 DR EMBL: AL117206; CAB60454.1; -  
 DR EMBL: AL110498; CAB60454.1; JOINED.  
 DR EMBL: AL110498; CAB57911.1; -  
 DR EMBL: AL117206; CAB57911.1; JOINED.  
 DR HSPD: P00736; IAPQ.  
 DR InterPro: IPR000152; ASX\_hydroxyl.  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR001881; EGF\_CA.  
 DR InterPro: IPR002049; Laminin\_EGF.  
 DR Pfam: PF00008; EGF; 25.  
 DR PRINTS: PR00011; EGF\_LAMININ.  
 DR SMART: SM00179; EGF\_CA; 4.  
 DR SMART: SM00001; EGF\_like; 18.  
 DR PROSITE: PS00010; ASX\_HYDROXYL; 4.  
 DR PROSITE: PS00022; EGF\_1; "NKR-WN\_22".  
 DR PROSITE: PS0136; EGF\_2; 24.  
 DR PROSITE: PS0187; EGF\_CA; 3.  
 DR Calcium-binding: EGF-like domain; Glycoprotein; Hydroxylation; Repeat.  
 KW SEQUENCE 1664 AA; 179279 MW; A64F093B4C7D5832 C9C64;

Query Match 12.38; Score 321; DB 5; Length 1664;  
 Best Local Similarity 34.48; Pred. No. 7.1e-22;  
 Matches 63; Conservative 22; Mismatches 66; Indels 32; Gaps 7.  
 QY 211 CEAAQKWGPCNHLCTACMNNGVCHEDTQECICPPGPGMGTCEKACELRTFGRTCKRCSG 270  
 DB 896 CHGPGGHIWMDLKKUNALDLSNLSFAPQWSKKTKWAAPITPKKESKKD- 447  
 QY 271 QGKSKSYVFLPLFVNSCATGKGLNEAHPSTVEFGKLEKLSNNLEMLHQP 328  
 DB 948 ---CAEAMHLPSTGELTPPKKKHKDETUSHLNFAFKKLSQUNATDSVTGSQ 1004  
 QY 330 LCPSPCMQGLQCEPHGIPRMTPIVCLPHIEVNSGKPE---NPTCKASGWLPTNEMTL 385  
 DB 1005 PCFPCWPKKKCLP-----PCPD-----GRFGHCNAICDCT-----TINDTSM 1042  
 QY 386 VKP 388  
 DB 1043 YNP 1045

RESULT 5  
 ID 088281 PRELIMINARY; PRT; 1574 AA.  
 AC 088281;  
 DT 01-NOV-1998 (TRIMBLrel. 08, Created)  
 DT 01-NOV-1998 (TRIMBLrel. 08, Last sequence update)  
 DT 01-DEC-2001 (TRIMBLrel. 19, Last annotation update)  
 DE MEGF6.  
 GN MEGF6.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=9836089; PubMed=9693030;  
 RA Nakayama M., Nakajima D., Kikuno K., Ohara O.  
 RT Identification of high-molecular-weight proteins with multiple EGF-like motifs by motif-trap screening.  
 RL Genomics 51:27-34 (1998).  
 DR EMBL: AB011552; BAA32452.1; -  
 DR HSPD: P00736; IAPQ.  
 DR InterPro: IPR000152; ASX\_hydroxyl.  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR001881; EGF\_CA.  
 DR Pfam: PF00008; EGF; 24.  
 DR SMART: SM00179; EGF\_CA; 4.  
 DR SMART: SM00001; EGF\_like; 19.  
 DR PROSITE: PS00010; ASX\_HYDROXYL; 5.

DR PROSITE: PS00022; EGF\_1; UNKNOWN\_23.  
 DR PROSITE: PS01186; EGF\_2; 23.  
 DR PROSITE: PS01187; EGF\_CA; 5.  
 KW Calcium-binding: EGF-like domain; Glycoprotein; Hydroxylation; Repeat.  
 SW SEQUENCE 1574 AA; 165445 MW; 2B48533DBF77F6E7 C9C64;  
 Query Match 12.28; Score 317.5; DB 11; Length 1574;  
 Best Local Similarity 35.08; Pred. No. 1.4e-21;  
 Matches 57; Conservative 21; Mismatches 62; Indels 24; Gaps 4.  
 QY 211 CEAAQKWGPCNHLCTACMNNGVCHEDTQECICPPGPGMGTCEKACELRTFGRTCKRCSG 270  
 DB 777 CPEGPGHGGCQFICPACHGASCPDETCTCLCPFCVCSRCQDTCSAGWCTGCTGCA 845  
 QY 271 QGKSKSYVFLPLFVNSCATGKGLNEAHPSTVEFGKLEKLSNNLEMLHQP 328  
 DB 836 ---FANDGHCPTTGTGSCAPGWTCLSCQPAQCHSGHWGPTTHPCNSAGHGNDAVSCI 892  
 QY 329 CLCSPCMQGLQCEPHGIPRMTPIVCLPHIEVNSGKPEPICK 371  
 DB 893 CLCEAYESPRQEQ-----SCQGYGFSCE 918

RESULT 6  
 ID 096KG7 PRELIMINARY; PRT; 1140 AA.  
 AC 096KG7;  
 DT 01-DEC-2001 (TRIMBLrel. 19, Created)  
 DT 01-DEC-2001 (TRIMBLrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TRIMBLrel. 19, Last annotation update)  
 DE MEGF10 PROTEIN (KIAA1780).  
 GN MEGF10.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=HIPPOCAMPUS;  
 RA Nagase T., Nakayama M., Nakajima D., Kikuno K., Ohara O.  
 RT Prediction of the coding sequences of unidentified human genes. XX.  
 RT The complete sequences of 100 new cDNA clones from brain which code  
 RT for large proteins in vitro.\*  
 RL DNA Res. 8:85-95 (2001).  
 DR EMBL: AB058676; BAB47409.1; -  
 SW SEQUENCE 1140 AA; 122204 MW; 45H2FA2J94J4895A C9C64;

Query Match 12.18; Score 316; DB 4; Length 1140;  
 Best Local Similarity 29.48; Pred. No. 1.3e-21;  
 Matches 67; Conservative 13; Mismatches 58; Indels 90; Gaps 6.  
 QY 159 PTSAFTRILVKKTAQAWGPECNHLCTACMNNGVCHEDTQECICPPGPGMGTCEKACELRTFGRTCKRCSG 271  
 DB 216 YTGAFCEDL---CPGKHSQCEQRCPQNGSVCHIVIGECSCPSGWMGTVAQGPCEG 271  
 QY 244 -----PFGMGTCEKAEELHTGTCKEFGSGEGPK 275  
 DB 272 PEGKNTSQVQTHNCTTAAAGGCHSGVYQWQIIVGLYGVFAELDGVNCK 331  
 QY 276 SY-----VFCLPDYQ-----CSQATGW 294  
 DB 332 CYHVSAGLCEAGFAGECEARLPEGLYGIKDKRCPCHLENTHSHPMSGECAKPGW 391  
 QY 284 KGIQCNPAQCHGEPDCKKLPSCNNGEMTIRFQGCLGSPQWGLQC 340  
 DB 302 SGIYNETSTFTYEAQACGQICSNADQLSVTSGTATAGFGAGIQC 439  
 RESULT 7  
 ID 096KG6

ID Q96K06 PRELIMINARY: PRT: 969 AA.  
 AC Q96K06;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DE MGFP11 PROTEIN (K1AA1781).  
 GN MGFP11.  
 OS Homo sapiens (Human).  
 OC Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,  
 OC Mammalia, Eutheria, Primates, Catarrhini, Hominoidea, Hominidae, Homo.  
 OX NCBI\_TaxID=9606.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=BRAIN;  
 RX MEDLINE:21245130, PubMed 11447406.  
 RA Nagase T., Nakayama M., Nakajima D., Kikuno K., Ohara O.;  
 RT "Prediction of the coding sequences of unidentified human genes. XX.  
 RT the complete sequences of 106 new cDNA clones from brain which code  
 RT for large proteins in vitro.";  
 RL DNA Res. 8:85-95(2001).  
 DR EMBL: AB058677; BAB47410.1; -.  
 SQ SEQUENCE 969 AA, 10165 MW, 55604.000 kDa, 9694.

Query Match 12.1%, Score 314, DB 4, Length 969,  
 Best Local Similarity 41.6%, Pred. No. 1.6e-21;  
 Matches 57; Conservative 14; Mismatches 52; Indels 14; Gaps 5;  
 QY 211 CEAKKWKPEFNHLCTAMNNVWVHEDTGFICPPFMPTREKATCELHFGPTKEP--C 268  
 DB 405 CPSTGWLGNESCT-CANCAACSPIDGSCCTPCWLDTCCLPCDGTGFLGNCSEHDC 463  
 QY 269 SQBQKSKYVFLPDY--GQSTATWKLQNEACHPGFGYGPDKLCKPNNEMDCRF 326  
 DB 464 SHADGC-----DPTVTHGQCLAGWTGIRDTCTPPGRWGPNGSVSCSENGSCSPE 515  
 QY 327 QC-CICSPGWQGIQPCR 342  
 DB 516 DGSCEAPGEGFLQUR 532

RESULT 8  
 Q14162  
 ID Q14162 PRELIMINARY: PRT: 830 AA.  
 AC Q14162;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
 DE KIAA0149 PROTEIN.  
 GN KIAA0149.  
 OS Homo sapiens (Human).  
 OC Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,  
 OC Mammalia, Eutheria, Primates, Catarrhini, Hominoidea, Hominidae, Homo.  
 OX NCBI\_TaxID=9606.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=96127530, PubMed-8590280;  
 RA Nagase T., Seki N., Tanaka A., Ishikawa K., Nomura N.;  
 RT "Prediction of the coding sequences of unidentified human genes. IV.  
 RT the coding sequences of 40 new genes (K1AA0121-K1AA0160) deduced by  
 RT analysis of cDNA clones from human cell line KG-1.";  
 RL DNA Res. 2:167-174(1995)  
 DR EMBL: D63483; BAA09770.1; -.  
 DR HSSP: F01140, 3B2.  
 DR InterPro: IPR000561; EGF-like.  
 DR SMART: SM00181; EGF; 5.  
 DR PROSITE: PS00022; EGF\_1; UNKNOWN\_6.  
 DR PROSITE: PS01186; EGF\_2; 6.  
 KW EGF-like domain; Glycoprotein.  
 SQ SEQUENCE 830 AA, 87431 MW, 81748.84 kDa, 9694.

Query Match 11.1%, Score 289, DB 4, Length 830;  
 Best Local Similarity 36.1%, Pred. No. 3.6e-19;  
 Matches 65; Conservative 14; Mismatches 79; Indels 22; Gaps 9;  
 QY 211 CEAKKWKPEFNHLCTAMNNVWVHEDTGFICPPFMPTREKATCELHFGPTKEP--C 270  
 DB 210 CPSTGWLGNESCT-CANCAACSPIDGSCCTPCWLDTCCLPCDGTGFLGNCSEHDC 266  
 QY 271 QESPCKSYVFLPDY--GQSTATWKLQNEACHPGFGYGPDKLCKPNNEMDCRF 328  
 DB 267 --KCKKNEAGSPDGSSESESESESESESESESESESESESESESESESESESESE 324  
 QY 329 -C-LCSPGWQGIQPCR-----REGIPRMTPKIVD-----LPDHFVNSGKFNPKICKAS 373  
 DB 325 HQGEPGLPWSHPPPEEHTFTSTFGECASTLFLVASSDTVTIRGVSAVWSPSCNAS 384  
 RESULT 9  
 Q43701  
 ID Q43701 PRELIMINARY: PRT: 830 AA.  
 AC Q43701;  
 DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
 DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE ACETYL LDL RECEPTOR PROTEIN.  
 OS Homo sapiens (Human).  
 OC Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,  
 OC Mammalia, Eutheria, Primates, Catarrhini, Hominoidea, Hominidae, Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=98058897, PubMed-9495444;  
 RA Adachi H., Tsujimoto M., Arai H., Inoue K.;  
 RT "Expression cloning of a novel scavenger receptor from human  
 RT endothelial cells.";  
 PL J. Biol. Chem. 272:31217-31220(1997).  
 DR EMBL: D86864; HAA24070.1; -.  
 DR HSSP: P01180; 2B2.  
 DR InterPro: IPR000561; EGF-like.  
 DR SMART: SM00181; EGF; 5.  
 DR PROSITE: PS00022; EGF\_1; UNKNOWN\_6.  
 DR PROSITE: PS01186; EGF\_2; 6.  
 KW EGF-like domain; Glycoprotein; Receptor; Signal.  
 FT SIGNAL 1 15 POTENTIAL.  
 FT CHAIN 16 830 ACETYL LDL RECEPTOR.  
 SQ SEQUENCE 830 AA, 87433 MW, 85604.000 kDa, 9694.

Query Match 11.1%, Score 289, DB 4, Length 830;  
 Best Local Similarity 36.1%, Pred. No. 3.6e-19;  
 Matches 65; Conservative 14; Mismatches 79; Indels 22; Gaps 9;  
 QY 211 CEAKKWKPEFNHLCTAMNNVWVHEDTGFICPPFMPTREKATCELHFGPTKEP--C 270  
 DB 210 CPSTGWLGNESCT-CANCAACSPIDGSCCTPCWLDTCCLPCDGTGFLGNCSEHDC 266  
 QY 271 QESPCKSYVFLPDY--GQSTATWKLQNEACHPGFGYGPDKLCKPNNEMDCRF 328  
 DB 267 --KCKKNEAGSPDGSSESESESESESESESESESESESESESESESESESESESE 324  
 QY 329 -C-LCSPGWQGIQPCR-----REGIPRMTPKIVD-----LPDHFVNSGKFNPKICKAS 373  
 DB 325 HQGEPGLPWSHPPPEEHTFTSTFGECASTLFLVASSDTVTIRGVSAVWSPSCNAS 384  
 RESULT 10  
 Q9Y151  
 ID Q9Y151 PRELIMINARY: PRT: 594 AA.  
 AC Q9Y151;  
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)

```

DE BC0NA:GH03529.
IN BC0NA:GH03529 OR CG2086.
OC Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Rubin G.M., Wan K.H., Harvey D., Lewis S.E., Brokstein P., Tsang G.,
RA Aquayani A., Arcina T.T., Baxter E., Blazej R.G., Butenhol C.,
RA Change M., Chavez C., Chew M., Doyle C.M., Farfan D.E., Frise E.,
RA Galle R., George R.A., Harris N.L., Hoskins R.A., Evans-Helm M.,
RA Houston K.A., Jammasti S.R., Kim E., Li P., Moshrefi M., Paeleeb J.M.,
RA Park S., Queiroz A., Sethi H., Snir E., Svirskas R.R., Weinburg T.,
RA Gelaiker S.E.
RA "Full length Drosophila melanogaster cDNA sequence".
RT Submitted (Apr-1999) to the EMBL/GenBank/DBJ databases.
RE EMBL: AF145617; AAD38592.1; -.
DR HSSP: Q12784; IHRF.
DR FlyBase: FBgn0027594; RefSeq: GH03529
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR002049; Laminin_EGF.
DR Pfam: PF00008; EGF_3.
DR PRINTS: PR00011; EGF_LAMININ.
DR SMART: SM00181; EGF_4.
DR PROSITE: PS00022; EGF_1; UNKNOWN_5.
DR PROSITE: PS01186; EGF_2; 5.
DR PROSITE: PS00290; IG_MHC; UNKNOWN_1
KW EGF-like domain; Glycoprotein.
SQ SEQUENCE 594 AA; 66234 MW; 1D6C95283E255014 CRC64;

Query Match 11.0%; Score 287.5; DB 5; Length 594;
Best Local Similarity 37.9%; Pred. No. 32c-19;
Matches 55; Conservative 16; Mismatches 61; Indels 13; Gaps 5;

QY 216 RCTEAKWKEPEVNHQVTA:MNNIV:HEIDRE:TP:IP:PM:PT:PKA:ELHTD:PT:KEP:~ 268
DB 118 KCLRGGVGGHAGIIC:R:LNNSSCDPSGNC:IC:SAWGTGALCAEP:PP:PY:GMECKERC 176
QY 269 ---SQEGV:ASVPE:PLP:PY:GSC:CA:GKWKGLQCN:ACH:PG:Y:G:K:K:K:SCN:GEM:CK 325
DB 177 ELLHGNKSCDHTIGDI:-----LCRTYIGITTEHP:PA:LY:PA:K:K:K:N:EH:GE:PH 240
QY 326 FQG:GLCSAGWGLOGERE:IPRMT 349
DB 231 VTGQCQCLPGWTGSNC:NESCPHT 254

RESULT 11
QY00A1 PRELIMINARY; PPT: 594 AA.
AC QY00A1
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE BC0NA:GH03529 PF07FIN
GN BC0NA:GH03529 OR CG2086.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Gelniker S.F., Holt P.A., Evans C.A., Goodyne T.D.,
RA Amanalides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.C., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

```

```

RA Brandon R.C., Rogers Y., H.C., Blazej R.G., Champ M., Pfeiffer R.D.,
RA Wan K.H., Doyle C., Baxter E.C., Helt C., Nelson C.R., Miklos G.L.G.,
RA Abill J.F., Aquayani A., An H.-J., Andrews-Piankoff C., Bailey D.,
RA Bailey P.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.V., Benes P.V., Bertram R.P., Bhandari D., Bolchakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Butlis K.C., Busam D.A., Butler H., Cadieu E., Cantor A., Chandra L.,
RA Cherry J.M., Casley S., Dahike C., Davenport L.R., Davies P.,
RA de Pablos B., Delcher A., Dena Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ikerow C.,
RA Jatali M., Kalush F., Karpen G.H., Ke Z., Keimison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasok P., Lei Y., Lavitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattel H., McIntosh T.C., Melced M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarri C., Morris J., Moshrefi A.,
RA Meunier S.M., Moy M., Murphy B., Murphy T., Murry D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paeleeb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders K.D.C., Schechter F., Shen H.,
RA Shue B.C., Siddiqui K., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Turner R., Venier E., Wang A.H., Wang X.,
RA Svirskas R., Tector C., Turner R., Venier E., Wang A.H., Wang X.,
RA Wang Z., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodard T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Ye F., Yeri J.S., Zhang W., Zhou X., Zhu S., Zhu X., Smith H.B.,
RA Gibbs R.A., Myers R.W., Rubin G.M., Venter J.C.
RA "The genome sequence of Drosophila melanogaster".
RL Science 287:2185-2195(2000).
DR EMBL: AE023472; AAF47552.1; -.
DR HSSP: Q12784; IHRF.
DR FlyBase: FBgn0027594; RefSeq: GH03529.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR002049; Laminin_EGF.
DR Pfam: PF00008; EGF_3.
DR PRINTS: PR00011; EGF_LAMININ.
DR SMART: SM00181; EGF_4.
DR PROSITE: PS00022; EGF_1; 5.
DR PROSITE: PS01186; EGF_2; 5.
DR PROSITE: PS00290; IG_MHC; UNKNOWN_1.
KW EGF-like domain; Glycoprotein.
SQ SEQUENCE 594 AA; 66212 MW; 8F1B5963F654579A CRC64;

Query Match 11.0%; Score 286.5; DB 5; Length 594;
Best Local Similarity 37.9%; Pred. No. 4e-19;
Matches 55; Conservative 16; Mismatches 61; Indels 13; Gaps 5;

QY 210 RCFACKWGFPCNHLCTACMNNNGVCHDPTGICHPGHMGPTCKACHLHPTGCTQERC~ 268
DB 118 KCLRGGVGGHAGIIC:R:LNNSSCDPSGNC:IC:SAWGTGALCAEP:PP:PY:GMECKERC 176
QY 269 ---SQEGV:ASVPE:PLP:PY:GSC:CA:GKWKGLQCN:ACH:PG:Y:G:K:K:K:SCN:GEM:CK 325
DB 177 ELLHGNKSCDHTIGDI:-----LCRTYIGITTEHP:PA:LY:PA:K:K:K:N:EH:GE:PH 240
QY 326 FQG:GLCSAGWGLOGERE:IPRMT 349
DB 231 VTGQCQCLPGWTGSNC:NESCPHT 254

RESULT 12
QY00A1 PRELIMINARY; PPT: 220 AA.
AC QY00A1
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)

```



```

RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=98360089; PubMed=9636030;
RA Nakayama M., Nishijima T., Nishise T., Nemura N., Saki N., Ohara O.;
RT "Identification of high-molecular-weight proteins with multiple EGF-
RT like motifs by motif-trap screening.";
RL Genomics 51:27-34(1998).
DR EMBL: AB011539; DAA42467.1; -
DR BSSP: P02468; IKD0;
DR InterPro: IPR000561; EGF-like;
DR InterPro: IPR002049; Laminin_EGF;
DR Pfam: PF00008; EGF; 2;
DR SMART: SM00180; EGF_Lam; 1;
DR PROSITE: PS00001; EGF_Like; 2;
DR PROSITE: PS00022; EGF_1; UNKNWN_3;
DR PROSITE: PS01186; EGF_2; 2;
KW EGF-like domain; Glycoprotein;
FT NON_TER 1
SQ SEQUENCE 153 AA; 15579 MW; FF6E34E53803F49D CRC64;

```

```

Query Match 9.78; Score 253.5; DB 4; Length 153;
Best Local Similarity 34.88; Pred. No. le-16;
Matches 57; Conservative 12; Mismatches 72; Indels 23; Gaps 6;

QY 223 LVTAMNNVCVHEFTGEYTPPEEMETGEKKAELHTEHGTGKEPSSQDECKSYVEGLF 282
DB 2 LCV-CUHGACDPISGRCLCFAGTGHGFCERGCPSGEGCQGRCD---CDGAGPCDP 56
QY 283 DNYGSCATGNGGCGNIACHGPGYGPDKKICSCNNGEMEDRFQG-CGCSPOWGGLQGE 341
DB 57 VTGLGLCFPRGSGAICNLKCFEGGQPSGLRIGVGGALDPSVSSQGRVGVGMPTC- 115
QY 342 REGIPRMTHKIVLDLPHIEVNSGRNP-ICKASGMPLPTNEEMT 384
DB 116 RRGGRRLRPE-----NPSLAGGSAGTLPASSRPT 144

```

```

RESULT 15
QYBOXD
ID Q9H0X0 PRELIMINARY; PRI; 117 AA.
AC Q9H0X0;
DE 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DI 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE PROTEIN-TYROSINE KINASE TIE2 (FRAGMENT).
OS Sus scrofa (Pig).
OC Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21153163; PubMed=11230987;
RA Kim J., Moon S.O., Han C.Y., Pak Y.K., Moon S.K., Kim J.J., Koh G.Y.;
RT "The angiopoietin-tie2 system in coronary artery endothelium prevents
RT oxidized low-density lipoprotein-induced apoptosis.";
RL Cardiovasc. Res. 49:872-881(2001).
DR EMBL: AF251494; AAK18753.1; -
DR InterPro: IPR003961; FN_111;
DR Pfam: PF00041; FN3; 1;
DR SMART: SM00060; FN3; 1;
KW Kinase;
FT NON_TER 1
FT NON_TER 117
SQ SEQUENCE 117 AA; 13161 MW; 6C86HE99BF40H6 CRC64;

```

```

Query Match 8.58; Score 221; DB 6; Length 117;
Best Local Similarity 100.0%; Pred. No. le-13;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 431 GMVEKPFNLSVKVLKPLNAPNVIDIGHNFAVINISSEPYFG 472

```

```

DB 1 GMVEKPFNLSVKVLKPLNAPNVIDIGHNFAVINISSEPYFG 42

Search completed: September 23, 2002, 22:16:47
Job time: 123 sec

```







## alignment\_scores:

Quality: 2505.00 Length: 472  
 Ratio: 5.519 Gaps: 0  
 Percent Similarity: 100.000 Percent Identity: 100.000

## alignment\_block:

US-09-733-764-2\_copy\_1\_472 x HUMTERPPTK

Align seq 1/1 to: HUMTERPPTK from: 1 to: 4138

1 MetAspSerLeuAlaSerLeuValLeuCysGlyValSerLeuLeuLeuSe 17  
 149 ATGAGATCTTTAGCTAGCTTGTGAGATCTTTGAGCTTGTGAGCTT 198  
 17 rGlyThrValGluGlyAlaMetAspLeuLeuLeuLeuLeuLeuLeu 34  
 199 TGGAACTGTGCAAGGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 248  
 34 euValSerAspAlaGluThrSerLeuThrCysLeuLeuLeuLeuLeu 50  
 249 TGTATCTGATGCTGCAACATCTGCTGCTGCTGCTGCTGCTGCTGCT 298  
 51 ProHisGluProLeuThrLeuGlyArgAspPheGluAlaLeuMetAsn 67  
 299 CCGATGAGCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 348  
 67 nHisGlnAspProLeuGluValThrGlnAspValThrArgGluTrpAla 84  
 349 GCACGAGATCCGCTGCAAGCTTCTCAAGATGTCACACAGAAATGGCT 398  
 84 ySLysValValTrpLysArgGluLysAlaSerLysLeuLeuAlaTyr 100  
 399 AAAAGTTGTTTGGAGAGAGAAAGGCTAGTAAAGTCAATGCTGCTTAT 448  
 101 PheCysGluGlyArgValArgGlyGluAlaLeuLeuGlyThrMetLe 117  
 449 TTCTGTGAAGGCTGCTTCTCAAGATGTCACACAGAAATGGCTGCT 498  
 117 sMetArgGlnGlnAlaSerPheLeuProAlaThrLeuThrMetThrVal 134  
 499 GATCGCTCAACAGCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTT 548  
 134 sPLysGlyAspAsnValAsnLeuSerPheLysLysValLeuLeuLys 150  
 549 ACAAGGAGATCAAGCTGCAACATCTTTTCAAAAGGATGATTAAGAA 598  
 151 GluAspAlaValLeuLeuLysArgGlySerPheLeuLeuLeuLeuLe 167  
 599 GAAGATGAGTATTTACAAAGATGCTTCTTCTTCTTCTTCTTCTTCT 648  
 167 qHisGluValProAspLeuGluValHisLeuLeuLeuLeuLeuLeu 184  
 649 CATGAGATGCTGATATCTGAGAGTACACCTGCTTCTGCTGCTGCTG 698  
 184 InAspAlaGlyValTrpSerAlaArgTyrIleGlyAsnLeuPheThr 200  
 699 AGGATGCTGGAGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 748  
 201 SerAlaPheThrArgLeuLeuValArgGlyCysGluAlaGlnLysTr 217  
 749 TGGAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 798  
 217 yProGluCysAsnHisLeuCysThrAlaCysMetAsnAsnGlyValCys 234  
 799 ACCTGAATGCAACCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 848  
 234 IsGluAspThrGlyLeuCysLeuLeuLeuLeuLeuLeuLeuLeuLeu 250  
 849 ATGAAGATGCTGGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 898  
 251 CysGluLysAlaCysGluLeuLeuLeuLeuLeuLeuLeuLeuLeuLe 267  
 899 TGTGAGATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 948

267 rGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 384  
 949 GTGAGTGGACACACAGGCACTGAGCTGCTGCTGCTGCTGCTGCTG 998  
 284 rGlyGlyCysSerCysAlaThrGlyTrpLysGlyLeuGlyCysAsnGlu 300  
 999 GTATGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1048  
 301 AlaCysHisProGlyPheTyGlyProAspCysLysLeuArgCysSer 317  
 1049 GATGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1098  
 317 sAsnAsnLysLysLysLysLysLysLysLysLysLysLysLysLys 334  
 1099 CAACAAATGGGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1148  
 334 LyTrpGlnGlyLeuGlnCysGluArgGlyGlyLeuProArgMetLeu 350  
 1149 GATGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1198  
 351 LysLeuValAspLeuProAspHisLeuGluValAsnSerGlyLysPhe 367  
 1199 AAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1248  
 367 nProLeuCysLysAlaSerGlyTrpProLeuProThrAsnGluLeu 384  
 1249 TCCCATTTGCAAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1298  
 384 hrLeuValLysProAspGlyThrValLeuHisProLysAspPheAsn 400  
 1299 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1348  
 401 ThrAspHisPheSerValAlaLeuPheThrLeuHisArgIleLeuPro 417  
 1349 ACGATATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1398  
 417 oAspSerGlyValTrpValCysSerValAsnThrValAlaGlyMetVal 434  
 1399 TGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1448  
 434 LuLysProPheAsnLeuSerValLysValLeuProLysProLeuAsn 450  
 1449 AAAAGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1498  
 451 ProAsnValLeuAspThrGlyHisAsnPheAlaValIleAsnLysSer 467  
 1499 GCAAAAGTGTATTAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1548  
 467 rGluProTyrPheGly 472  
 1549 TGAACCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1564

seq\_name: qb\_cm:BTIE2A

seq\_documentation\_block:

LOCUS BTIE2A 4625 bp mRNA linear MAM 04-NOV-1993  
 DEFINITION B.taurus Tie 2 mRNA.  
 ACCESSION X71424  
 VERSION X71424.1 GI:296577  
 KEYWORDS receptor tyrosine kinase.  
 SOURCE cow.

## ORGANISM:

Bos taurus  
 Eukaryota; Metazoa; Chordata; Craniata, Vertebrata, Euteleostomi;  
 Mammalia, Eutheria, Cetartiodactyla, Ruminantia, Pecora, Bovidae;  
 Bovidae; Bovinae; Bos.

PEPENCE 1 (bases 1 to 4625)

AUTHORS Sato, T.N.

TITLE Direct Submission

JOURNAL Submitted (18-MAR-1993) T.N. Sato, Roche Institute of Molecular

REFERENCE 2 (bases 1 to 4625)

AUTHORS Sato, T.N., Gub.Y., Kozak, G.A. and Audus, K.L.

TITLE Pic-1 and tie-2 define another class of putative receptor tyrosine kinase genes expressed in early embryonic vascular system  
JOURNAL Proc. Natl Acad Sci U S A 56 (25) 9355-9358 (1993)  
MEDLINE 94022374  
REMARK Erratum: [[published erratum appears in Proc Natl Acad Sci U S A 1993 Dec 15;90(24):12056]]

FEATURES  
Source 1..4625

Location/Qualifiers  
1..4625

organism="Bos taurus"

db\_xref="taxon:9913"

cell\_type="endothelial cells"

tissue\_type="brain"

seq\_peptide 324..377

gene "tie 2"

gene "tie 2"

codon\_start=1

product="receptor tyrosine kinase"

protein\_id="CAA50555.1"

db\_xref="GI:2966578"

db\_xref="SWISS-PROT:Q06807"

1 tasatid tdm -MDLALVLA:GVSLLSATVGGAGLILINSLPLVSTAEISLT  
1 ASQWRPHEPTITIGRDPEALMNQHDPLEVTQDTRRWAKKVVVKREKASKINGATFC  
EORVGGQAIRITMTVMKMDQASFIIPATITMTVDRGNVNI SFKKVLKEEDAVIYNKCS  
FTHSPREHVPDILLEVQVHPAQPDAGVVSARY IGGNLTSAFTRLITVRQAQWGP  
ECNRCTACMNNHITCHEDTGECIPPGFMGTCEKAC EPHTEFGTCERKSEPEPSK  
FVGLDPPVAGLSATIGKGLQCNATJPGYGGDCKLRGCTNGEKCDREDFGLCSPG  
RQGLQTEKEVWPRMTKIEDI.PDHIEVNSGFKNP ICKASGWPPPAEEMTLVKPDDTV  
LRPKDNNHCHLSVAFTL INKILPPDSCVWVCVNIVSQMVEKPPNISKVLPKPLNA  
PKVIDTCHNPVAVINSEPYEGGPIKSKKLLIKVNVNHYEAWRHIOVTNEIVTLNLYE  
PRTEYECGLVRRGEGRGHGPVRRFTTASIGLPPRGISLILPKSOTTLNLTWOPI  
PESEDDPYEVEPSPVMSNDQNIKVPGLNLSVILNHLHPPEQYIVPAPVNTKAQG  
EWSEDLAWTSDVPPPENIKIFNITSSAVISWTILDGYSISAIIRYKVGQKNE  
DOHDIVKIKNATITQYKGLLEPQTVYQVDIFAENNTGSSNPTSSHELTLSESOAPA  
DIGRKMILLALIGSAGMIGLTVLLAPLIMLOIKRANVORHMAOAFONVREPPAVQFN  
SCTIALNKAANNRPIILPYVLWMDJLKQDVAQCNQFQV KAKIKKIDELMDAALK  
PMKLYASKDKNRPPAGHLYVLCGLGHHPNIIILGACERHGYLYLALFAPGNLLIDF  
LRKSRLEIDPAPFAINASTLSSQGLHFAALVARGMAYLSKGFTHRGLAARNLL  
GVENVAFVDFPISQFVYVVRKTMGPVPMWMALESYNYSVTITNSDWSYGVLLW  
EIVSLGTPYAGMTCAELEYKLFQYPLEKPLNIDQVYDLMHJWREKPYERPPSFAQ  
11VSLNRMLEKRYVNTILYKPYIYAGIIDCSAFEA"

gene 324..3701

gene="tie 2"

378..3698

gene="tie 2"

product="receptor tyrosine kinase"

BASE COUNT 1292 a 1064 c 1108 g 1162 t

ORIGIN

alignment\_scores:

Quality: 2445.00 Length: 472

Ratio: 5.259 Gaps: 0

Percent Similarity: 98.093 Percent Identity: 93.220

alignment\_block:

US-09-733-764-2\_COPY\_1\_472 a BT1E2A

Align seq 1/1 to: BT1E2A from: 1 to: 4625

1 MetAspSerLeuAlaSerLeuValIleuGlyValSerLeuLeuLeuSe 17

|||||

424 ATGATCTCTTATAGCGGCTTATGTTCTCTCGGAGTCAGTTGCTCTTC 373

|||||

17 rGlyThrValGluGlyAlaMetAspLeuLeuLeuLeuLeuLeuLeu 34

|||||

374 TGAACAGTGGACGGTCCATGCCAGCTGATGTCATGATCCCTACCTC 423

|||||

34 euValSerAspAlaGluThrSerLeuThrCysLeuAlaSerGlyTrpArg 50

|||||

424 TTGATCTGATGCTGAGACATCCCTACATCTGATCTGAGATGCGCC 473

|||||

51 ProHisGluProIleThrIleGlyAlaGlyAspPheGluAlaIleuMetAsnGlu 67  
|||||

474 GATCATATATATATATATATATATATATATATATATATATATATATATAT 523  
|||||

67 nHisGluAspProLeuGluValThrGlnAspValThrArgGluTrpAlaLeu 84  
|||||

524 GCACTGCAATATATATATATATATATATATATATATATATATATATATATAT 573  
|||||

84 ystLysValValIleTrpLysAspGluLeuAlaSerLysIleAsnGlyAlaTrp 100  
|||||

574 AAAAGTGTGTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 623  
|||||

101 PheCysLeuGlyArgValArgGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 117  
|||||

624 TTTTGTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 673  
|||||

117 smetArgGlnGlnAlaSerPheLeuProAlaThrLeuThrMetThrValAla 134  
|||||

674 CATGGGTCAACAAGCTGCTTCTTACACAGTACCTTAACTATGACATGCTGG 723  
|||||

134 splysGlyAspAsnValAsnIleSerPheLysLysValLeuLeuLysGlu 150  
|||||

724 AAGAGAGATATATGTCAGATATATATATATATATATATATATATATATATAT 773  
|||||

151 GluAspAlaValIleCysLysAsnGlySerPheIleHisSerValProArg 167  
|||||

774 GAAAGCTATATATATATATATATATATATATATATATATATATATATATAT 923  
|||||

167 qHisGluValProAspIleLeuGluValHisLeuProHisAlaGlnProG 184  
|||||

824 GATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 873  
|||||

184 leAspAlaGlyValTrpSerAlaArgTyrIleGlyValAsnLeuSerPheTrp 200  
|||||

874 AGATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 923  
|||||

201 SerAlaPheThrArgGluIleValArgCysGluAlaGlnLysIlePro 217  
|||||

924 TAAATATATATATATATATATATATATATATATATATATATATATATATATAT 973  
|||||

217 yProGluCysAsnIleLeuGlyThrAlaCysMetAsnGlyValCysH 234  
|||||

974 ACCTGAATGTAACCGCATCTGCACCTCTTGTATGAACAATGGCACTGCC 1023  
|||||

234 isclAspThrGlyGlyCysIleCysProGlyPheMetGlyArgThr 250  
|||||

1024 ATGAGAGATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1073  
|||||

251 CysCulysAlaCysGlnIleHisThrPheGlyArgThrCysLysLeuArg 267  
|||||

1074 TGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1123  
|||||

267 yCysSerGlyValGluGlyLysSerTrpValPheCysLeuTrpAspP 284  
|||||

1124 GTGTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1173  
|||||

284 rGlyGlyCysSerCysAlaThrGlyTrpLysGlyLeuGlnCysGlnGlu 400  
|||||

1174 CCTATGGGTGTTCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1223  
|||||

301 AlaCysHisProGlyPheTrpGlyProAspCysLysLeuArgCysSerCys 317  
|||||

1224 GATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1273  
|||||

317 sAsnAsnGlyGluMetCysAspArgPheGlnGlyCysLeuCysSerProG 344  
|||||

1274 CACCAATGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1424  
|||||

334 lyTrpGlnGlyLeuGlnCysGluArgGluGlyIleProArgMetThrPro 450  
|||||

1324 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1474  
|||||

351 LysIleValAspLeuProAspHisIleGluValAsnSerGlyLysPheAs 467  
|||||

```

1374 AACATAGAGATTTGAGATATATAGAGTAAAGTGTAAATTCAA 1423
367 nprotilecysylsalaserGlyThrProLeuProThrAspGluMetT 384
1424 CCCCATCTGTAATATATATATATATATATATATATATATATAT 1473
384 hrLeuValLysProAspGlyThrValLeuHisProLysAspPheAsnHis 400
1474 CCGTGGTCAAGCCAGATCGGACAGTGGTACGTCGCAAGAGATTTAAGCAC 1523
401 ThrAspHisPheSerValAlaAlaPheThrThrIleHisArgIleLeuProPr 417
1524 ATGGTAT 1573
417 oAspSerGlyValIrrpValCysSerValAsnThrValAlaGlyMetValG 434
1574 TCACTCAGCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1623
434 LuLysProPheAsnIleSerValLysValLeuProLysProLeuAsnAla 450
1624 AAAAGCCCTTCAACATCTGTTAAAGTCTTCCAAAACCCCTTGAATGCC 1673
451 ProAsnValIleAspThrGlyHisAspPheAlaValIleAsnIleSerSe 467
1674 CCAAGGTAT 1723
467 rGluProTyrPheGly 472
1724 TGAGCCTTACTTTGGG 1739

```

seq\_name: qb\_ro:MMTIE2A

seq\_documentation\_block:

```

LOCUS MMTIE2A 3490 bp mRNA linear ROD 04-NOV-1993
DEFINITION M.musculus Tie 2 mRNA.
ACCESSION X71426
VERSION X71426 1 GI:296612
KEYWORDS receptor tyrosine kinase.
SOURCE house mouse.
ORGANISM Mus musculus.
REFERENCE Eukaryota, Metazoa, Chordata, Craniata, Vertebrata, Euteleostomi,
AUTHORS Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 3490)
JOURNAL Sato, T.N.
REFERENCE Submitted (18-MAR-1993) T.N. Sato, Roche Institute of Molecular
AUTHORS Biology, 340 Kingsland St., Nutley, NJ 07035, USA
TITLE 2 (bases 1 to 3490)
REFERENCE Sato, T.N., Qin, Y., Kozak, C.A. and Andrus, K.L.
TITLE Tie-1 and tie-2 define another class of putative receptor tyrosine
JOURNAL kinase genes expressed in early embryonic vasculature: System
MEDLINE Proc Natl Acad Sci U S A 90 (1993) 9458-9458 (1993)
REMARK 94022374
Erratum: [published erratum appears in Proc Natl Acad Sci U S A
1993 Dec 15;90(24):12056]]
FEATURES
SOURCE 1..3490
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="4"
/tissue_type="lung"
/dev_stage="adult (6-8 weeks)"
sig_peptide 87..140
CDS 87..3455
/gene="Tie 2"
/codon_start=1
/product="receptor tyrosine kinase"
/protein_id="CAA50557.1"
/db_xref="GI:296613"
/db_xref="M00-M01-08664"
/db_xref="SWISS-PROT:Q02858"

```

```

/translation="MDSLAGLVAGSVLLYGVGAMDLILINSLPLVDAETSLTC
IASGWHPEPTTIGRDFEALMNHQDFLEVTQDTFEWAKKVVWREKASIKNGAYFC
EGRVQCAIRIRIKRMKQASHLPALLIMIVKQUNVNIKNVLIKEIDAVLYKNGS
FTHSVPRHEVPDILLEVHLPACQDAGVYSARYIGSNLFTSAFTRLIWKRCQAKWGP
DCSRPTCTCKNNGVCHEDFCGICPPQMGRTCEKACQPHUPGRICKKCSGPGCCS
YVCLLEFYTSATSWPGLNHEATTSYVGRDCKLKHHEEETLREGGGLLSQS
WGLQCEKGRGRTPTQIEDPDHLEVNSGFENPICKASGMPLPTSEEMTLVKPDGTV
LQPDENYTDRESVAIFTVNRVLPDSDGVWGSVNTVAGMVEKPENISVKVLPPLHA
PNVIDTGHNAFTINISSEYFGTPIKSKKLFYKPVNGAWKYIEVNIPTINYLEPK
IDYELACVOLARIGGCGHPGVRRFTTASIGLPPHKGSLILPKSOIALNUTWOPIFT
NSHDEYVFEVRSQGTSDOONIKVPONIISVLLSNIVPRPOYVPRPVNTKQCEW
SEELPAWILSDILPPPEPNIKISNIISLHMSWTIVDGYSSIIIRYKVGKNEQQ
HIDVKTNATVTOYQLGLEPTEYTHVDIEFNIGSSNPAPFSUELRTPHSPASADL
GAGKMLLIALLSAGMTCTIVLLAELIMQLKRVNQKMAAFUNKEEPVAVNSGT
LAINRKKNPNPTIYVLDMDIKFDVIGRNPQGVILKARIKKDGLRMDAAIKRKK
EYASKDHRDFAGLELVLCIKGHENIINLACAPHRGYLYLALEYAHGNLIJLKK
SVLETDPAFAIANSTASTLSQQLHFAADVARGMDYLSOKOPIHRLDAARNLIJGK
NYLAKTADGLSRGQEVYVKTMGBLPVRMALESINYSVYTTNSDQWSYGVLLWEIV
SLGTFYCGMTCAELYEKLQPYRLEKPIINCTIDEVLDLMPQCWREKPYRSPFAGILV
SLNRMLEERKTYVNTTLXYEKFTYAGIDCSAEAA"
87..3455
/gene="Tie 2"
141..3452
/gene="Tie 2"
/product="receptor tyrosine kinase"
BASE COUNT 964 a 811 c 896 g 819 t
ORIGIN

```

gene

mat\_peptide

/product="receptor tyrosine kinase"

BASE COUNT 964 a 811 c 896 g 819 t

ORIGIN

alignment\_scores:

Quality: 2407.00 Length: 472

Ratio: 5.233 Gaps: 0

Percent Similarity: 97.458 Percent Identity: 91.949

alignment\_block:

US-09-733-764-2\_copy\_1\_472 x MMTIE2A

Align seq 1/1 to MMTIE2A from 1 to 3490

```

1 MetAspSerLeuAlaSerLeuValIleuCysGlyValSerLeuLeuLeuSe 17
|||||
87 ATGATATCTTTAT 136
|||||
17 rGlyThrValGluGlyAlaMetAspLeuLeuLeuLeuLeuLeuLeuLeuLeu 34
|||||
137 TGGAGTAGTAAAGCTGCCATGGACCTGATCTTGAICAAATTCCTACCTC 186
|||||
34 euValSerAspAlaGluThrSerLeuThrCysIleAlaSerGlyTyrArq 50
|||||
187 TTGTCTCTATGCGGAAATATCTTATCTATATATATATATATATATATATATAT 236
|||||
51 PROHISGLPTTITTEHRIIEGLYARGASPPHEGLUALALEUMETASNGL 67
|||||
237 TGGAT 286
|||||
67 ohisGlnAspProLeuGluValThrGlnAspValThrArgGluTrpAlaL 84
|||||
287 GCACCAAGATCCACGAGAGATTATTCAGAGATGTGACACAGCAATGGCGGA 336
|||||
84 yLysValValIrrpLysArgGluLysAlaSerLysIleAsnGlyAlaTyr 100
|||||
337 AAAAAGTTGTTTGGAAAGAGAAAAAGGCAATAGATTAATTAATTAATTAATTAAT 386
|||||
101 rheCysGluGlyArgValArgGlyClnAlaAlaIleCargIleArgThrMetIly 117
|||||
387 TTCTGTGAGATGTGAGATTCTGAGGATAGGATATAAGATATAAGGATATAGAA 436
|||||
117 sMetArgGlnGlnAlaSerPheLeuProAlaThrLeuThrMetThrVala 134
|||||
437 GATGGCTCAACAAAGCTCCTTCCTACCTGCTACCTTTAATTAATTAATTAATTAAT 486
|||||
134 sPLysGlyArgAsnValAsnIleSerPheLysLysValLeuIleLysGlu 150
|||||

```

```

487 ACAAGGAGATAATGTGAACATATCTTTCAAAAAGGTGTAAATTAAGAA 536
151 GluAspAlaValIleValTrpLysAsnGlySerPheIleHisSerValProArg 167
|||||
547 CAACATCGACGATGATACAAAAAGCGCTCTTCAICCACTCAGTCCGCCG 586
|||||
167 qHisGluValProAspIleLeuGluValHisLeuProHisAlaGluProG 184
|||||
587 GCATCAAGTACCTCATATTTACAAGTTCACATTCGGCGCATGCTCAGCCGC 636
|||||
184 InuAspAlaGlyValTrpSerAlaSerArgTyrIleCysIleAsnLeuPheThr 200
|||||
647 AGATGCTGTCTGTCTATCTGGTCAAGATATATAGAGAAATTTGTTTAA 686
|||||
201 SerAlaPheThrArgIleuIleValArgArgCysGluAlaCysIleTyrP 217
|||||
687 TCACGCTTCACCGCCGATCTCTGGACATGCAATGCAACATTCAGCTGGCG 746
|||||
217 TrpGluCysAsnAlaIleuLysThrAlaCysMetAsnAsnGlyValCysH 234
|||||
747 GCGGACATG:AGCGGCGCTGTGATCTACTTCCACAGAACAAATCGAGTCTGC 786
|||||
244 IscIAspThrGlyCysIleCysIleCysIleCysIleCysIleCysIleCys 250
|||||
787 ATGAGATACGGAGGAAGATCTGGCGTCTCTGATTTATATGAGAAATTA 846
|||||
251 CysIleLysAlaCysIleCysIleCysIleCysIleCysIleCysIleCys 267
|||||
847 TGTCAAGAACTCTTCAGCGCCACACATTTGCCAGCACCTCTAAAGAAAG 886
|||||
267 ArgSerCysIleCysIleCysIleCysIleCysIleCysIleCysIleCys 284
|||||
887 GGTGATGAGTAAATCAAGATCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 946
|||||
284 rArgGlyCysSerCysAlaIleGlyTrpLysGlyLeuGluCysAsnGlu 300
|||||
947 CTTAGCGGTATTCTGTGCTGATCAAGTCAAGTCAAGTCAAGTCAAGTCA 986
|||||
301 AlaCysHisProGlyPheThrGlyTrpAspCysLysLysLysLysLysLys 317
|||||
987 GCATGCCCATCTCTCTACTACAGCACAGCACAGTCTGTAAGCTCAGTGCAC 1036
|||||
417 AsnAspGlyLeuMetCysAspArgPheCysIleCysLeuCysSerPro 334
|||||
1037 TACCAATCAACAGATATGATGATCGTTCAAGCAATGCCCTCTCTCTCAAG 1086
|||||
334 TrpTrpCysGlyLeuGluCysGlyArgGlyCysIleCysIleCysIleCys 350
|||||
1087 GATGGTAAGGAGTGTGATCTGTAAGAAAGTCAAGTCAAGTCAAGTCAAG 1136
|||||
451 LysIleValAspLeuProAspHisIleGluValAsnSerGlyLysPheAs 367
|||||
1137 CAGATAGAGATTTGGCAGATACATTTGAAGTAAACAGTGGAAATTTAA 1186
|||||
367 nProIleCysLysAlaSerGlyTrpProLeuProThrAsnGluGluMet 384
|||||
1187 GCGCATCTGTAAGCGCTCTGGGCGGCACTACTACTAGTGAAGAAATCA 1236
|||||
384 TrpLeuValLysProAspGlyIleValLeuHisProLysAspPheAsnHis 400
|||||
1247 CCGTATGAGATTAATAGAGATATGATGATTAATTAATTAATTAATTA 1286
|||||
401 ThrAspHisPheSerValAlaIlePheThrIleHisArgIleLeuPro 417
|||||
1287 ACACATCTCTCTCTAGTATATATATATATATATATATATATATATAT 1346
|||||
417 oAspSerGlyValTrpValCysSerValAsnThrValAlaGlyMetValG 434
|||||
1337 TGACTCAGGAGTCTGGTCTCTGAGTCAAGTCAAGTCAAGTCAAGTCAAG 1386
|||||
434 LysLysTrpPheAsnIleSerValLysValLeuProLysProLysProLys 450
|||||
1487 AAAAGGCTTTCAATATTTCAGTAAAGATCTTCAAGAGAGAGAGAGAGAG 1436

```

```

451 ProAsnValIleAspThrGlyHisAsnPheAlaValIleAsnIleSerSe 467
|||||
1437 CCAATCTGATTCACACTGCACATAAATCTTGTATCATCAATACAGCTC 1486
|||||
467 rGluProTyrPheGly 472
|||||
1487 TGAGCCTTACTTTGGG 1502
|||||
seq_name: gb_pat:E08401
seq_documentation_block:
Locus E08401
DEFINITION cDNA of mouse TIE-2 receptor.
ACCESSION F08401
VERSION F08401.1 GI:2176518
KEYWORDS JP 1994315382-A/1.
SOURCE Mus sp.
ORGANISM Mus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE
1 (bases 1 to 3760)
AUTHORS Murayama,Y., Suda,T., Iwama,A., Yasunata,K. and Masuyasu,Y.
TITLE TIE-2 RECEPTOR AND DNA CODING THE SAME RECEPTOR
JOURNAL Patient: JP 1994415482-A 1 15-Nov 1994;
COMMENT:
CS Mus sp. (mouse)
PN JP 1994315382-A/1
PD 15-NOV-1994
PE 06-MAY-1994 JP 1993129912
FI: MURAYAMA YU-YO, SUDA TOSHIO, IWAMA ATSUSHI, YASUNATA KUNIO, FI
MASUYASU YASUHIKO
PI: FJ215712, F07814700, C12121702;
CC strandedness: Double;
CC topology: linear;
CC Feature is identified by similarity;
FH key Location/Qualifiers
FH source 1..3760
/organism="Mus sp."
/cell_line="DA-1"
FH seq_peptide 152..211
FH rat_peptide 212..3523
/product="mouse TIE-2 receptor".
FEATURES
source location/Qualifiers
1..3760
/organism="Mus sp."
/db_xref="taxon:10095"
BASE COUNT 1024 a 876 c 970 q 890 t
ORIGIN
alignment_scores:
Quality: 2407.00 Length: 472
Ratio: 5.233 Gaps: 0
Percent Similarity: 97.458 Percent Identity: 91.94%
alignment_block:
US-09-733-764-2_copy_1_472 x E08401
Align seq 1/1 to: E08401 from: 1 to: 3760
1 MetAspSerLeuAlaSerLeuValIleuCysGlyValSerLeuLeuLeuSe 17
|||||
152 AlaAsnValIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuI 201
|||||
17 rGlyThrValGluGlyAlaMetAspLeuLeuIleuLeuAsnSerLeuProL 44
|||||
202 GGTACATATAGAAAGGAGGATATGCAAGTCAATGATGATGATGATGATG 251
|||||
34 cuValSerAspAlaGluThrSerLeuThrCysIleAlaSerGlyTrpArg 50
|||||
252 TTGTGTCTATGTCGAAAATATGATGATGATGATGATGATGATGATGATG 301

```







174 TGGAGTACTAGAAAGGAGATGAGATTTTATATTAATTCCTTACCTC 223  
 34 euValSerAspAlaGluThrSerLeuThrCysIleAlaSerGlyTrpArg 50  
 224 TTGTCCTGTGATGGCAAAACATCCCTCACCCTGCATTGCCCTCGGGTGGCAC 273  
 51 ProHisGluProIleThrIleGlyArgAspPheGluAlaLeuMetAsnGly 67  
 274 CCCATCAGCCCATCATCACATAGCAAGGACCTTTCAAGCCCTTAATGAACCA 323  
 67 nHisGlnAspProLeuGluValThrGlnAspValThrArgGluTrpAlaIle 84  
 324 GCACCAAGATCCATGGAGGTTATCTAAGATGTAGCAGAGAAATGGGCGA 373  
 84 yslIysValValTrpIysArgGluIysAlaSerIysIleAsnGlyAlaIle 100  
 374 AAAAGTCTCTTGGCAAGACAGCAAAAGGCCACTTAAGATTAAATGGTCTAT 423  
 101 PheCysGluGlyArgValArgGlyGluAlaIleArgIleArgThrMetIle 117  
 424 TTCTGTGAAGTTCAGTTCGAGTTCGACAGGCTATAAGGATAACGACATGAA 473  
 117 smcArgGlnGlnAlaSerPheLeuProAlaThrLeuThrMetThrValAla 134  
 474 GATGGGTCAAAAGCATCTCTCTGCTTACTTGTACTTTAACTATGACCTGG 523  
 134 splIysGlyAspAsnValAsnIleSerPheIysValLeuIleIysGlu 150  
 524 ACAGGGGAGATAATCTGAACAATATCTTTTCAAAAAGGCTTTAAATTAAAGAA 573  
 151 GluAspAlaValIleTyrLysAsnGlySerPheIleHisSerValProAr 167  
 574 CAAGAATGAGTCAATTTACAAAATGGCTCCCTCAATCCACTCAGTCCCGCG 623  
 167 gHisGluValProAspIleLeuGluValHisIleuProHisAlaGlnProG 184  
 624 GCATGAATACCTGATATTTTAAAGTTCAATTGAGGATATGTCAGGAGCC 673  
 184 InAspAlaGlyValIysSerAaargIyrIleGlyGlyAsnLeuPheThr 200  
 674 AGGATCTGGTGTGACCTCGGCGAGGTACATAGGAGGAAACCTGTTTACCC 723  
 201 SerAlaPheThrArgLeuIleValArgCysGluAlaGlnLysTrpGly 217  
 724 TCAGCTTCACCAAGCTGATGTTCCGAGCATCTGAAGCTCAGAACTGGGG 773  
 217 yProGluCysAsnHisIleCysThrAlaCysMetAsnAsnGlyValCysH 234  
 774 GCGGACTGTATACCTGCTGATATTTTAAATTTATTTTAAATATATTTG 823  
 234 IscIleAspThrGlyGluCysIleCysProTrpGlyPheMetGlyArgThr 250  
 824 ATGAAGATACCGGGAAATCCATTTGCCCTCTGGGTTTATGGGAGAGCA 873  
 251 CysGluLysAlaCysGluLeuHisThrPheGlyArgThrCysLysGluAr 267  
 874 TGTGAGAAAGCTTGGTAAATGACACATTTGCGACAGGATCTTAAAGAAAG 924  
 267 gCysSerGlyGlnGluGlyCysIysSerTyrValPheCysLeuProAspP 284  
 924 GTGATATATTAATAAAGGATGTAATTTATTTATTTTATTTTATTTTAT 974  
 284 rCysGlyCysSerCysAlaThrGlyTrpIysGlyLeuGlyCysAsnGly 300  
 974 CTACGCTGCTCTGCTGCGACAGGCTGGAGGGGTTCACAGTGGCAAGCA 1023  
 301 AlaCysHisProGlyPheTyrGlyProAspCysLysLeuArgCysSerC 317  
 1024 GCATGCCCATCTGGTCTATACAGGACACACACATCTAAGCTCAGGTGCCACTG 1073  
 317 sAsnAsnGlyGluMetCysAspArgPheGlnCysIleuCysSerProG 334  
 1074 TACTCAATGAAGATATGTGATGTGATGTGATGTGATGTGATGTGATGTG 1123

334 lyTrpGlnGlyLeuGlnCysGluArgGlyGlyIleProArgMetThrPro 350  
 1124 GATGGCAAGGGCTGGTATGCTGATCAAGAAAGGCAAGGCTCAAGCATCAGCCCA 1173  
 351 LysIleValAspLeuProAspHisIleGluValAsnSerGlyLysPheAs 367  
 1174 CAGATAGAGATTGCTAGATTAATTAAGATTAATTAAGATTAATTAAGATTA 1223  
 367 nProIleCysLysAlaSerGlyTrpProLeuProThrAsnGluGluMetP 384  
 1224 CCCATCTCTTAAGATCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 1273  
 384 hrLeuValLysProAspGlyThrValLeuHisProLysAspPheAsnHis 400  
 1274 CCCAGGAAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAG 1323  
 401 ThrAspHisPheSerValAlaIlePheThrIleHisArgIleLeuProPr 417  
 1324 ACAGATGCTTCTTATGCTGCTTATATTAATTAATTAATTAATTAATTAAT 1373  
 417 oAspSerGlyValTrpValCysSerValAsnThrValAlaGlyMetValG 434  
 1374 TGACTAGAGATCTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1423  
 434 luLysProPheAsnIleSerValLysValLeuProLysProLysProLysAla 450  
 1424 AAAAGGCTTTCACATTTCCGTCACAAAGTCTTCCAGAGGCGCTGCCAGGC 1473  
 451 ProAsnValIleAspThrGlyHisAsnPheAlaValIleAsnIleSerSe 467  
 1474 CCAATCTGATTAAGATGAGATTAATTAATTAATTAATTAATTAATTAAT 1523  
 467 rGluProTyrPheGly 472  
 1524 TAAATTTATTTTGG 1539  
 seq\_name: qb\_pat:171113  
 seq\_documentation\_block:  
 LOCUS 171113  
 DEFINITION Sequence 1 from patent US 5681714.  
 ACCESSION 171113  
 VERSION 171113.1 GI:3007248  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 4175)  
 AUTHORS Breitman, M.L. deceased, Rossant, J., Dumont, D.J. and Yamaquuchi, T.P.  
 TITLE Nucleic acid encoding tek receptor tyrosine kinase  
 JOURNAL Patent: US 5681714-A 1 28-OCT-1997;  
 FEATURES  
 Location/Qualifiers  
 source  
 1..4175  
 /organism="unknown"  
 BASE COUNT 1175 a 942 c 1047 g 1011 t  
 ORIGIN

alignment\_scores:  
 Quality: 407.09 Length: 472  
 Ratio: 5.233 Gaps: 0  
 Percent Similarity: 97.458 Percent identity: 91.949  
 alignment\_block:  
 US-09-733-764-2\_COPY\_1\_472 x 171113  
 Align seq 1/1 to: 171113 from: 1 to: 4175

1 MetAspSerLeuAlaSerLeuValLeuCysGlyValSerLeuLeuSe 17  
 124 ATGATATTTTAAAGGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTT 173



[illegible]

seq documentation block:

ACCESSION I71115

VERSION 171113.1 GI:3007Z30  
KEYWORDS

SOURCE Unknown.

ORGANISM  
UNKNOWN.  
Unclassified

PFPPFNFCE 1 (bases 1 to 4176)

**ACTIONS** Richardson, M. L. (1984). **TITLE** *Neotoma and Neotoma*

JOURNAL, Patent: US 5681714-A

| FEATURES | LOCATION/QUANTITY |
|----------|-------------------|
| 1        | 1176              |

/organism-

BASE COUNT 11/6 a 939 c  
ORIGIN

argument=20000.

Ratio: 5,233

PERCENT SIMILARITY: 97.438 F

alignment\_block:

US-09-133-164-2\_COPY\_1\_4/2 X 17

Align seq 1/1 to: I71115 from

```

1 MetAspSerLeuAlaSerLeuValLeuCysGlyValSerLeuLeuLeuSe 17
|||||
124 ATGGACTCTTAAAGCGGTTATTTCTGTGGAGTCAGCTGCTCTCTTA 173
17 rGlyThrValGluGlyAlaMetAspLeuLeuLeuLeuLeuLeuLeu 34
|||||
174 TGAGTATGACAAAGCGGTAAGAGTCACTGCTGATCAATTCCTCACTC 223
34 euValSerAspAlaGluThrSerLeuThrCysIleAlaSerGlyTrpArg 50
|||||
224 TTGTGCTGTGATGCGGAACATGCTGACCTGCTGCTGCTGCTGCTGCT 273
51 ProHisGluProIleThrIleGlyArgAspPheGluAlaLeuMetAsnG 67
|||||
274 GCGCATGAGCGCATACATAGAGAGAGAGAGAGAGAGAGAGAGAGAG 323
67 nHisGluAspProLeuGluValThrGlnAspValThrArgGluTrpAla 84
|||||
424 GATCAAGATCCACATGCGAGCTTACACAGATGTCACAGAGAAATGGCG 373
84 yLysValValTrpLysArgGluLysAlaSerLysIleAsnGlyAlaTrp 100
|||||
374 AAAATGTTGTGGCAAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 423
101 PheCysGluGlyArgValArgGlyGluAlaIleArgIleArgThrMetLy 117
|||||
424 TTCTGTGAAGTGTGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 473
117 sMetArgGluGlnAlaSerPheLeuProAlaThrLeuThrMetThrVal 134
|||||
474 GATGCGGTACAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 523
134 sPlyGlyAspAsnValAsnIleSerPheLysLysValLeuIleLysGlu 150
|||||
524 ACAGGCGAGATAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 573
151 GluAspAlaValIleTyrLysAsnGlySerPheIleHisSerValProAr 167
|||||
574 GAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 623
167 gHisGluValProAspIleLeuGluValHisLeuProHisAlaGlnProG 184
|||||
624 GCAATGAGATGCTGATATTTAAGAGTACATGCTGCTGCTGCTGCTG 673
184 InAspAlaGlyValTyrSerAlaArgTyrIleGlyGlyAsnLeuPheThr 200
|||||
674 AGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 723
201 SerAlaPheThrArgLeuIleValArgArgCysGluAlaGlnLysTrpG 217
|||||
724 TCAGCTTCAAGAGTGTATTTGAGATGTTGAGATGTTGAGATGTTGAG 773
217 yProGluCysAsnHisLeuCysThrAlaCysMetAsnGlyValCysH 234
|||||
774 GCGCGATGTAAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 823
234 IsdAlaPheThrGlyGlyCysIleCysProProGlyPheMetGlyArgThr 250
|||||
824 ATCAACATATATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 873
251 CysGlyLysAlaCysGluLeuLeuSerPheGlyArgThrCysLysGlyAr 267
|||||
874 TGAGAGAAAGCTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 923
267 qYsserGlyGluGluGlyCysLysSerTyrValPheCysLeuProAspP 284
|||||
924 GTATATGAG 973
284 rTyrThrLysSerAlaSerCysIleGlyGlyGlyGlyGlyGlyGlyGly 300
|||||
974 CTACAG 1023

```

```

301 AlaCysHisProGlyPheThrGlyProAspCysIleCysLeuGlyCysSerCy 417
|||||
1024 GCATGAGATATCTGGTAACTAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1073
317 sAsnAsnGlyGluMetCysAspArgPheGlnGlyCysLeuCysSerProG 434
|||||
1074 TACCAATCAAGACATATATGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1123
334 LyTrpGlnGlyLeuGluCysGluArgGluGlyIleProArgMetThrPro 450
|||||
1124 GATGATGAGAGATGATGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1173
351 LysIleValAspLeuProAspHisIleGluValAsnSerGlyLysPheAs 467
|||||
1174 CAAATAGAGATTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1223
367 nProLysLysAlaSerGlyTrpProLeuProThrAsnGluGluMet 484
|||||
1224 GCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1273
384 hLeuValLysProAspGlyThrValLeuHisProLysAspPheAsnHis 400
|||||
1274 GCTATATGAG 443
401 ThrAspHisPheSerValAlaIlePheThrIleHisArgIleLeuProP 417
|||||
1324 ACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1373
417 oAspSerGlyValTrpValCysSerValAsnThrValAlaGlyMetValG 434
|||||
1374 TGATCAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1423
434 LuLysProPheAsnIleSerValLysValLeuProLysProLeuAsnAla 450
|||||
1424 AAAAGCTTTCAGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1473
451 ProAsnValIleAspIleHisAsnPheAlaValIleAsnIleSerSer 467
|||||
1474 GCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1523
467 rGluProTyrPheGly 472
|||||
1524 TCAGCTTACTTTGG 1539

```

seq\_name: gb\_to:MMTEK

seq\_documentation\_block:

LOCUS MMTEK 4176 bp mRNA linear R00 04-DEC-1993

DEFINITION M. musculus mRNA for tek.

ACCESSION X67553.340311

VERSION X67553.3 GI:297158

KEYWORDS tek gene; Tyrosine kinase.

SOURCE house mouse.

ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 2587 to 4176)

Dumont, P.J., Yamaqucha, J.P., Gordon, R.A., Rossant, J., and

Breitman, M.L.

tek, a novel tyrosine kinase gene located on mouse chromosome 4, is expressed in endothelial cells and their presumptive precursors

Oncogene 7 (8), 1471-1480 (1992)

9234855

2 (bases 1 to 4176)

Dumont, P.J.

Direct Submission

Submitted (26-APR-1993) D.J. Dumont, Samuel Lunenfeld Res.

Institute, Mount Sinai Hospital, 600 University Ave., Toronto,

Ontario, M5G 1X5, CANADA

3 (bases 1 to 4176)

Dumont, D.J., Gradwohl, G.J., Fong, G.H., Auerbach, R., and

Breitman, M.L.

The endothelial-specific receptor tyrosine kinase, tek, is a member















[illegible][illegible]

1. The first part of the document discusses the importance of maintaining accurate records of all transactions and activities. It emphasizes that proper record-keeping is essential for transparency and accountability, particularly in financial matters. The text outlines various methods for organizing and storing data, including digital databases and physical filing systems. It also mentions the need for regular audits and reviews to ensure the integrity of the information.

2. The second part of the document focuses on the role of communication in achieving organizational goals. It highlights the importance of clear and concise communication, both internally and externally. The text provides guidelines for effective communication, such as using appropriate language, listening actively, and providing feedback. It also discusses the benefits of open communication and how it can foster a collaborative and innovative work environment.

3. The third part of the document addresses the challenges of managing resources and personnel. It discusses the importance of efficient resource allocation and the need for a skilled and motivated workforce. The text provides strategies for managing personnel, including recruitment, training, and performance management. It also mentions the importance of maintaining a positive work culture and providing opportunities for professional development.

4. The fourth part of the document discusses the importance of risk management and contingency planning. It emphasizes that organizations should always be prepared for unexpected events and should have a plan in place to minimize the impact of such events. The text outlines various risk management techniques, such as risk assessment, risk mitigation, and risk transfer. It also mentions the importance of regular risk reviews and updates to the contingency plan.

5. The fifth part of the document discusses the importance of innovation and continuous improvement. It emphasizes that organizations should always be looking for ways to improve their processes and products. The text provides guidelines for fostering innovation, such as encouraging creativity, providing resources for research and development, and implementing a culture of continuous improvement. It also mentions the importance of staying up-to-date with the latest trends and technologies in the industry.

OM c1: us-09-733-764-2\_copy\_1\_472.rng N\_Geneseq\_032802 \* out\_format pfs  
 Date: Sep 23, 2002 11:25 PM  
 About: Results were produced by the GeneCore software, version 4.5.  
 Copyright (c) 1993-2000 CompuGen Ltd.

## Command line parameters:

-MODEL=frame\_12n\_model -DEV=xlh  
 -O=/cdm2\_1/USP02-SP-01/0209733764/09042002/20020924/very\_fast\_1\_545  
 -DH=N\_Geneseq\_032802 -CPMT=fastap -SUFFIX=ring -CAPOP=12\_000  
 -GAPEXT=4\_000 -MINMATCH=0 -LUN=1000000 -LUNEXP=0\_000  
 -QCAPOP=4\_500 -GAPEXT=0\_050 -XCAPOP=10\_000 -XGAPEXT=0\_500  
 -FCAPOP=6\_000 -FPAPEXT=7\_000 -YCAPOP=10\_000 -YGAPEXT=0\_500  
 -DELOP=6\_000 -FPAPEXT=7\_000 -STAPT=1 -MATRIX=blotsum62  
 -TRANS=human40 rdi -T1T1=45 -XCAP=1000000 -THP=SCORP-pct  
 -THK MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs  
 -NCRM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
 -USER=US09733764\_RCGNL\_1\_0 -NCPUP=6 -ICPU=3 -LONGLOG  
 -DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -NO\_XLPXY -WAIT -THRAOS=1

## Search information block:

Query: US-09-733-764-2\_COPY\_1\_472  
 Query length: 472  
 Database: N\_Geneseq\_032802:  
 Database sequences: 1736436  
 Database length: 858457221  
 Search time (sec): 221.93000

## score\_list:

| Sequence                                                   | Strd Orig | ZScore  | RScore   | len | Documentation |
|------------------------------------------------------------|-----------|---------|----------|-----|---------------|
| /SIDSL/gcdata/geneseq/geneseq-emb1/NA1994.DAT:AAQ55179 +   | 2605.00   | 3317.35 | 2.0e-176 | 41  |               |
| /SIDSL/gcdata/geneseq/geneseq-emb1/NA1995.DAT:AAQ55334 +   | 2407.00   | 3064.64 | 2.4e-162 | 37  |               |
| /SIDSL/gcdata/geneseq/geneseq-emb1/NA1994.DAT:AAQ62129 +   | 2407.00   | 3063.68 | 2.7e-162 | 41  |               |
| /SIDSL/gcdata/geneseq/geneseq-emb1/NA2000.DAT:AA240648 +   | 2407.00   | 3063.68 | 2.7e-162 | 41  |               |
| /SIDSL/gcdata/geneseq/geneseq-emb1/NA2000.DAT:AA240640 +   | 2407.00   | 3063.67 | 2.7e-162 | 41  |               |
| /SIDSL/gcdata/geneseq/geneseq-emb1/NA1994.DAT:AAQ43740 +   | 2407.00   | 3063.27 | 2.7e-162 | 41  |               |
| /SIDSL/gcdata/geneseq/geneseq-emb1/NA1995.DAT:AAQ51999 +   | 2407.00   | 3062.71 | 3.0e-162 | 41  |               |
| /SIDSL/gcdata/geneseq/geneseq-emb1/NA1994.DAT:AAQ56697 +   | 2494.00   | 3045.74 | 2.7e-161 | 41  |               |
| /SIDSL/gcdata/geneseq/geneseq-emb1/NA1993.DAT:AAQ47064 +   | 865.50    | 1090.16 | 2.2e-52  | 38  |               |
| /SIDSL/gcdata/geneseq/geneseq-emb1/NA2001B.DAT:AA579371 +  | 688.50    | 859.02  | 1.7e-49  | 6   |               |
| /SIDSL/gcdata/geneseq/geneseq-emb1/NA1993.DAT:AAQ47065 +   | 675.50    | 847.14  | 7.7e-39  | 37  |               |
| /SIDSL/gcdata/geneseq/geneseq-emb1/NA2001B.DAT:AAH19879 +  | 331.00    | 415.43  | 8.9e-15  | 4   |               |
| /SIDSL/gcdata/geneseq/geneseq-emb1/NA2001B.DAT:AAH19878 +  | 331.00    | 406.48  | 2.7e-14  | 3   |               |
| /SIDSL/gcdata/geneseq/geneseq-emb1/NA2001B.DAT:AAH19874 +  | 331.00    | 391.73  | 1.8e-13  | 1   |               |
| /SIDSL/gcdata/geneseq/geneseq-emb1/NA2001B.DAT:AAH19872 +  | 318.50    | 398.44  | 7.5e-14  | 14  |               |
| /SIDSL/gcdata/geneseq/geneseq-emb1/NA2001A.DAT:AAQ99715 +  | 318.50    | 398.44  | 7.5e-14  | 1   |               |
| /SIDSL/gcdata/geneseq/geneseq-emb1/NA2001A.DAT:AAQ99760 +  | 318.50    | 397.47  | 8.6e-14  | 1   |               |
| /SIDSL/gcdata/geneseq/geneseq-emb1/NA2000.DAT:AAZ61678 +   | 318.50    | 397.43  | 8.7e-14  | 16  |               |
| /SIDSL/gcdata/geneseq/geneseq-emb1/NA2001A.DAT:AAQ59961 +  | 318.50    | 397.43  | 8.7e-14  | 1   |               |
| /SIDSL/gcdata/geneseq/geneseq-emb1/NA2001A.DAT:AAQ52807 +  | 315.00    | 395.74  | 1.1e-13  | 1   |               |
| /SIDSL/gcdata/geneseq/geneseq-emb1/NA2001A.DAT:AAH34884 +  | 315.00    | 391.24  | 1.9e-13  | 3   |               |
| /SIDSL/gcdata/geneseq/geneseq-emb1/NA2001B.DAT:AAQ588013 + | 313.50    | 390.35  | 2.1e-13  | 1   |               |
| /SIDSL/gcdata/geneseq/geneseq-emb1/NA2001B.DAT:AAQ588013 + | 313.50    | 385.11  | 4.2e-13  | 2   |               |
| /SIDSL/gcdata/geneseq/geneseq-emb1/NA2001B.DAT:AAQ57219 +  | 312.50    | 394.85  | 1.3e-13  | 9   |               |
| /SIDSL/gcdata/geneseq/geneseq-emb1/NA2001A.DAT:AAQ27790 +  | 309.00    | 386.11  | 3.7e-13  | 1   |               |
| /SIDSL/gcdata/geneseq/geneseq-emb1/NA2001A.DAT:AAQ27789 +  | 309.00    | 381.12  | 7.0e-13  | 2   |               |
| /SIDSL/gcdata/geneseq/geneseq-emb1/NA2001A.DAT:AAQ27810 +  | 309.00    | 380.67  | 7.4e-13  | 2   |               |
| /SIDSL/gcdata/geneseq/geneseq-emb1/NA1996.DAT:AAQ19059 +   | 308.00    | 385.09  | 4.2e-13  | 14  |               |
| /SIDSL/gcdata/geneseq/geneseq-emb1/NA2001A.DAT:AAQ27792 +  | 307.00    | 381.29  | 6.9e-13  | 1   |               |
| /SIDSL/gcdata/geneseq/geneseq-emb1/NA2001A.DAT:AAQ27791 +  | 307.00    | 375.55  | 1.4e-12  | 3   |               |
| /SIDSL/gcdata/geneseq/geneseq-emb1/NA2001A.DAT:AAQ27788 +  | 300.50    | 368.47  | 3.6e-12  | 1   |               |
| /SIDSL/gcdata/geneseq/geneseq-emb1/NA2001A.DAT:AAQ27787 +  | 300.50    | 364.07  | 6.2e-12  | 1   |               |
| /SIDSL/gcdata/geneseq/geneseq-emb1/NA2001A.DAT:AAQ20811 +  | 300.00    | 373.45  | 1.9e-12  | 1   |               |
| /SIDSL/gcdata/geneseq/geneseq-emb1/NA2001B.DAT:AAQ198119 + | 299.00    | 369.16  | 3.2e-12  | 5   |               |
| /SIDSL/gcdata/geneseq/geneseq-emb1/NA2000.DAT:AAQ51078 +   | 298.00    | 362.30  | 2.9e-11  | 3   |               |
| /SIDSL/gcdata/geneseq/geneseq-emb1/NA2001A.DAT:AAQ27469 +  | 297.50    | 361.36  | 3.5e-11  | 3   |               |
| /SIDSL/gcdata/geneseq/geneseq-emb1/NA2001B.DAT:AAQ519826 + | 284.50    | 357.73  | 1.4e-11  | 1   |               |
| /SIDSL/gcdata/geneseq/geneseq-emb1/NA2001B.DAT:AAQ57362 +  | 281.50    | 354.07  | 2.2e-11  | 1   |               |
| /SIDSL/gcdata/geneseq/geneseq-emb1/NA2001B.DAT:AAQ586742 + | 277.00    | 357.41  | 1.5e-11  | 3   |               |

/SIDSL/gcdata/geneseq/geneseq-emb1/NA2001B.DAT:AAQ519826 + 277.00 357.41 1.5e-11  
 /SIDSL/gcdata/geneseq/geneseq-emb1/NA2002.DAT:AAQ4598 + 277.00 340.04 1.4e-10  
 /SIDSL/gcdata/geneseq/geneseq-emb1/NA2001A.DAT:AAQ27488 + 277.00 337.43 1.9e-10  
 /SIDSL/gcdata/geneseq/geneseq-emb1/NA2001A.DAT:AAQ27486 + 277.00 447.40 1.9e-10  
 /SIDSL/gcdata/geneseq/geneseq-emb1/NA2001B.DAT:AAQ586746 + 272.50 339.92 1.4e-10

seq\_name: /SIDSL/gcdata/geneseq/geneseq-emb1/NA1994.DAT:AAQ55179

seq\_documentation\_block:

US AAQ55179 Standard; cDNA to mRNA; 4138 BP.

XX AAQ55179;

ET 25-JUN-1994 (first entry)

XX Human orphan receptor kinase gene.

XX Ork gene, ligands, antibodies, PCR, amplification: ss.

OS Homo sapiens.

XX key location/Qualifiers

ET CDS 149..3523

ET misc\_feature 2744..3219

ET misc\_feature 149..3520

ET misc\_feature 203..3520

ET misc\_feature 149..3520

ET misc\_feature 203..3520

ET misc\_feature 149..3520

ET misc\_feature 203..3520

ET misc\_feature 149..3520

ET misc\_feature 203..3520

ET misc\_feature 149..3520

ET misc\_feature 203..3520

ET misc\_feature 149..3520

ET misc\_feature 203..3520

ET misc\_feature 149..3520

ET misc\_feature 203..3520

ET misc\_feature 149..3520

ET misc\_feature 203..3520

ET misc\_feature 149..3520

ET misc\_feature 203..3520

ET misc\_feature 149..3520

ET misc\_feature 203..3520

ET misc\_feature 149..3520

ET misc\_feature 203..3520

ET misc\_feature 149..3520

ET misc\_feature 203..3520

ET misc\_feature 149..3520

ET misc\_feature 203..3520

ET misc\_feature 149..3520

ET misc\_feature 203..3520

ET misc\_feature 149..3520

ET misc\_feature 203..3520

ET misc\_feature 149..3520

ET misc\_feature 203..3520

ET misc\_feature 149..3520

ET misc\_feature 203..3520

ET misc\_feature 149..3520

ET misc\_feature 203..3520

ET misc\_feature 149..3520

ET misc\_feature 203..3520

ET misc\_feature 149..3520

ET misc\_feature 203..3520

ET misc\_feature 149..3520

ET misc\_feature 203..3520

ET misc\_feature 149..3520

ET misc\_feature 203..3520

ET misc\_feature 149..3520

ET misc\_feature 203..3520

ET misc\_feature 149..3520

ET misc\_feature 203..3520

DNA and protein sequences for orphan receptor tyrosine kinase -  
 and expression vectors for production of recombinant protein and  
 antibodies specific for the protein, useful in research  
 Claim 1, Fig 1, 57pp; English.  
 Degenerate oligonucleotide primers based on the sequence suggested  
 in the kinase domain of all receptor tyrosine kinases was used for  
 PCR of single stranded cDNA from human placental polyA mRNA. PCR  
 prod. HKK-6 contained a novel sequence which was used as a probe to  
 isolate longer fragments from a human placental cDNA library. One  
 clone (sequence shown) contained the entire coding region and was  
 called the ork gene. The ork gene can be used as a search tool  
 in in vitro assays for detection of ork, its ligands or their  
 interactions.

Sequence 4138 BP, 1170 A, 911 C, 987 G, 1070 T, 0 other;

```

alignment_scores:
 Quality: 2605.00 length: 472
 Ratio: 5.519 Gaps: 0
 Percent Similarity: 100.000 Percent identity: 100.000

alignment_block:
us-09-733-764-2_copy_1_472 x AAQ55179 ..
Align seq 1/1 to: AAQ55179 from: 1 to: 4138

1 MetAspSerLeuAlaSerLeuValLeuGlyValSerLeuLeuLeuSe 17
149 ATGGACCTTTTAACTTAACTTAACTTAACTTAACTTAACTTAACTT 198
17 rGlyThrVal.GluGlyAlaMetAspLeuLeuLeuLeuLeuLeu 34
199 TGGACACTGGGAAGTGGATGATGATGATGATGATGATGATGATGAT 248
34 euValSerAspAlaGluThrSerLeuThrCysIleAlaSerGlyTrpArq 50
249 TTGTATTCATGCTGAAACATCTCAGCTGATTCGCTTCGCTGCGC 298
51 ProHisGluProIleThrIleGlyArqAspPheGluAlaLeuMetAsnG 67
299 GCGCATGAGGATCATGATGATGATGATGATGATGATGATGATGATG 348
67 nHisGlnAspProLeuGluValThrGlnAspValThrArgGluTrpAlaI 84
349 GACACAGATGATGATGATGATGATGATGATGATGATGATGATGATG 398
84 ySLysValValTrpLysArgGluLysAlaSerLysIleAsnGlyAlaTyr 100
399 AAAAACTGCTTTGGAAGACAGACAAAGCTAGTAAATCAATTCGCTTAT 448
101 PheCysGluGlyArqValArgGlyGluAlaIleArgIleArqThrMetLy 117
449 TTCTTTGAGAGGGCGATTCGAGGAGAGGATGAGGATGAGGATGAGG 498
117 sMetArgGlnGlnAlaSerPheLeuProAlaThrLeuThrMetThrVal 134
499 GATGCTTAACTAACTAACTAACTAACTAACTAACTAACTAACTAACT 548
134 sPLysGlyAspAsnValAsnIleSerPheLysValLeuLeuLysGlu 150
549 ACAAGGAGATACCGTGAATATCTTTCAAAAGGATGATGATGATGAT 598
151 GluAspAlaValIleTyrLysAsnGlySerPheIleHisSerValProAr 167
599 GAAGATGAGTGAATTA:AAAAATGATGATGATGATGATGATGATGAT 648
167 qHisGluValProAspIleLeuGluValHisLeuProHisAlaGlnProG 184
649 GATATAGCTAGCTGATATATGAGAGAGAGAGAGAGAGAGAGAGAGAG 698
184 InAspAlaGlyValTyrSerAlaArqTyrIleGlyAsnLeuPheThr 200
699 AGGATGCTGGAGTGAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 748
201 SerAlaPheThrArqLeuIleValArqArqCysGluAlaGlnLysTrpGl 217
749 TGGGCTTAACTAACTAACTAACTAACTAACTAACTAACTAACTAACT 798
217 PheCysGlyAsnHisLeuGlySerAlaCysMetAspAsnGlyValCysH 234
799 ACCTGATATGAACTAACTAACTAACTAACTAACTAACTAACTAACTAA 848
234 IsGluAspThrGlyGluCysIleCysProGlyPheMetGlyArqThr 250
849 ATCAACATATGAGCAATGATGATGATGATGATGATGATGATGATGAT 898
251 CysGluLysAlaCysGluLeuHisThrPheGlyArqThrCysLysGluAr 267

```

```

899 TGTGAGAGAGCTTGTGAACCTGCACACGCTTTCGACAGAACTTTAAAGAAAG 948
267 gCysSerGlyGlnGluGlyCysLysSerTyrValPheCysAlaThrAspP 284
949 GTGACGTGGATAAATAGCAAGTAAAGTCTTAAGGCTGCTGCTGCTGCTGCT 998
284 rCysGlyCysSerCysAlaThrGlyTrpLysGlyLeuGlnCysAsnGlu 300
999 CCTATGGGCTTCTCTGTCGACACAGCTGCAAGGGCTGCTGCTGCTGCTGCT 1048
301 AlaCysHisPheGlyPheTyrClyPheAspCysLysLeuArqCysSerC 317
1049 GATATGCAATGCTGCTTCTTAAAGGCAATATTAAGTAAATTAAGTAAAG 1098
317 sAsnAsnGlyGluMetCysAspArqPheGlnGlyCysLeuCysSerProG 334
1099 CAACATGGGAGACATGCTGATGCTTCCAAAGGATGCTCTGCTGCTGCTGCT 1148
334 LyTrpGlnGlyLeuGlnCysGluArqGlnGlyIlePheArgMetThrPro 350
1149 GATGTAAGGGGCTTCAGCTGCAATAGAAAGGCTATACCGAGGATCACCGCA 1198
351 LysIleValAspLeuProAspHisIleGluValAsnSerGlyLysPheAs 367
1199 AAGATAGTGGATTTGCTATATATATATATATATATATATATATATATAT 1248
367 nProIleCysLysAlaSerGlyTrpProLeuProIleThrAsnGluMet 384
1249 TCTTATTTCTAAAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1298
384 hrLeuValLysProAspGlyThrValLeuLysProLysAspPheAsnHis 400
1299 CCTGCTGTAAGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1348
401 ThrAspHisPheSerValAlaIlePheIleThrIleHisArgIleLeuPro 417
1349 AGGATCATTTCTATATATATATATATATATATATATATATATATATAT 1398
417 oAspSerGlyValTrpValCysSerValAsnThrValAlaGlyMetValG 434
1399 TGACCTCAGGAGTTTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1448
434 LuLysProPheAsnIleSerValLysValLeuProLysProLeuAsnAla 450
1449 AAAAGCCCTTCACATTTCTGCTTAAAGCTGCTGCTGCTGCTGCTGCTGCT 1498
451 ProAsnValIleAspThrGlyHisAsnPheAlaValIleAsnIleSerSe 467
1499 CCAAACGCTGATTCACCTGGACATAAATCTTCTGCTGCTGCTGCTGCTGCT 1548
467 rGluProLysPheGly 472
1549 TGAGCCTTACTTTGGG 1564

seq_name: /SIDS1/gcdata/geneseq/geneseq-emb1/NA1995.DAT:AAQ75334
seq_documentation_block:
XX AAQ75334 standard: cDNA; 3760 BP.
XX AC
XX XX
XX DT 24-Apr-1995 (first entry)
XX DE
XX XX Marine tyrosine kinase receptor tic 2 cDNA.
XX XX Marine tyrosine kinase receptor; tic-2; hematosis diagnosis; ds.
XX OS Mus musculus.
XX XX
XX XX Location/Qualifiers
XX FT CDS 152..4523
XX FT /*tag= a
XX FT sig_peptide 152..211

```

FT mat\_peptide /\*tag- b  
 FT 212...3520 /\*tag- c  
 XX JP06315382-A.  
 XX 15-NOV-1994.  
 XX 06-MAY-1994: 93JP-0129912.  
 XX 06-MAY-1993: 93JP-0129912.  
 XX (SUDA/) SUDA T.  
 XX (YAMA ) YAMANOUCHI PHARM CO LTD.  
 XX WP1: 1995-032331/05.  
 XX P-PSDB: AAR67391.  
 DR DNA coding a tie-2 receptor and a tie 2 receptor used in the  
 XX diagnosing hematosis  
 PS Claim 2, Pages 13-17, 17pp, Japanese.  
 XX AA075334 encodes AAR67391 the murine tyrosine kinase receptor tie-2.  
 CC The cDNA is thought to participate in hematosis, and can therefore  
 CC be used in the diagnosis of hematosis.  
 XX Sequence 3760 BP: 1024 A: 876 C: 970 G: 890 T: 0 other:  
 SQ  
 alignment\_scores:  
 Quality: 2407.00 Length: 472  
 Ratio: 5.233 Gaps: 0  
 Percent Similarity: 97.458 Percent Identity: 91.949  
 alignment\_block:  
 US-09-733-764-2\_copy\_1\_472 x AA075334  
 Align seq 1/1 to: AA075334 from: 1 to: 3760  
 1 MetAspSerLeuAlaSerLeuValLeuCysGlyValSerLeuLeuLeuSe 17  
 152 ATGGACTCTTTAGCGGCTTAGTTCCTGCGGAGGACAGCTTGGCTTTA 201  
 17 rGlyThrValGlyAlaMetAspLeuLeuLeuLeuLeuLeuLeuProL 34  
 202 TGAGATAGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 251  
 34 euValSerAspAlaGluThrSerLeuThrCysIleAlaSerGlyTrpArg 50  
 252 TTGTGTGTGATGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 301  
 51 ProHisGluProIleThrIleGlyArgAspPheGluAlaLeuMetAsnG 67  
 302 CCCCATGAGCCATCAGCATAGGAAGGAGCTTTGAAGCCCTTAATGAACCA 351  
 67 nHisGluAspProLeuGluValThrGlnAspValThrArgGluTrpAlaL 84  
 352 GCACCAAGATCTCTGGAGAGATTACTAAGATGTAAAGAAATGGAGGA 401  
 84 ystysValValTrpIlyArgGluIysAlaSerIlystIleAsnGlyAlaTy 100  
 402 AAAAAAGTTGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 451  
 101 PheCysGluGlyArgValArgGlyGluAlaIleArgIleArgThrMetLy 117  
 452 TTCTGTGAAGCTCCAGTTCCAGGACAGGCTATAAGCATACCGACCAAGAA 501  
 117 sMetArgGlnGlnAlaSerPheLeuProAlaThrLeuThrMetThrVal 134  
 502 GATGGGTAAAGAGGCTCTTCTTAACTGATACCTTTAATCATGACCTGG 551  
 134 splysGlyAspAsnValAsnIleSerPheIlystysValLeuIleCysGlu 150

552 ACAGGACATATAAAGTCAAAATATACTTCAAAAAGGCTGTTAAATAAGAA 601  
 151 GluAspAlaValIleTyrIlySerAsnGlySerPheIleIleIleSerValProAr 167  
 602 GAAGATGAGAGTATTTATATAAAATGAGTCTCTCACTCACTCACTCACTCA 651  
 167 gHisGluValIProAspIleGluValHisLeuProHisAlaGlnProG 184  
 652 GGATGAAGTACCTGATATTTAGAAAGTCTCTTGAAGTCTCTCTCTCTCT 701  
 184 InAspAlaGlyValIlySerAlaArgTyrIleGlyClyAsnLeuPheThr 200  
 702 AGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 751  
 201 SerAlaPheThrArgLeuIleValArgCysGluAlaGlnLeuIlyTrpG 217  
 752 TCAGCTCTTCACTAGGTTCATTTGAGATATGAGAAATTAAGAGAGAGAG 801  
 217 yProGluCysAsnHisLeuCysThrAlaCysMetAsnAsnGlyValCysH 234  
 802 GAGGAAATGATAGAGGCTCTTGTATTAATTTAAAGAAATTAAGAGAGTCT 851  
 234 isClnAspThrGlyGluCysIleCysIlySerIlyrValIleCysLeuProAsp 250  
 852 ATCAAGATACCGGGGAAATGCAATTTGCCCTCTCTGCTGCTGCTGCTGCTG 901  
 251 CysGluIlyAlaCysGlnLeuHisThrPheGlyArgThrCysLysGluAr 267  
 902 TGTGAGAAAGCTTGTGAGCCGCCACACATTTGGCAGGAGCTGTAAAGAAAG 951  
 267 gCysSerGlyGlnGluClyCysLysSerIlyrValIleCysLeuProAsp 284  
 952 GTGTAGTGGAGAGAGAGAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1001  
 284 rGlyThrCysSerCysAlaThrGlyTyrIlyGlyClyLeuGlnCysAsnGlu 300  
 1002 CTTCAGGGGCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1051  
 301 AlaCysHisProGlyPheTyrGlyProAspCysLysLeuArgCysSerCy 317  
 1052 GCATGCTCATCTGCTTACTACGAGCAGACATGTAAGCTCAGCTGCTGCTGCT 1101  
 317 sAsnAsnGlyGluMetCysAspArgPheGlnGlyCysLeuCysSerProG 334  
 1102 TACCAATGAAGAGATATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1151  
 334 IlyTrpGlnGlyLeuGlnCysGluArgGluGlyIleProArgMetThrPro 350  
 1152 GATGGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1201  
 351 LysIleValAspLeuProAspHisIleGluValAsnSerGlyLysPheAs 367  
 1202 CATATAGAGAGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1251  
 367 nProIleCysIlyAlaSerGlyTrpProLeuProThrAsnGluGluMetL 384  
 1252 CTTCTCTCTGTAAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1301  
 384 hLeuValIlyProAspGlyThrValLeuHisProIlyAspPheAsnHis 400  
 1302 CCTACTGAAGCCAGATGGGACAGTCTCCACCAACCAATCATTCAACTAT 1351  
 401 ThrAspHisPheSerValAlaIlePheThrIleHisArgIleLeuProPr 417  
 1352 ATCATGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1401  
 417 oAspSerGlyValTrpValCysSerValAsnThrValAlaGlyMetValG 434  
 1402 TCACTCAAGAGCTGAGTCTTGAATGAGAGAGAGAGAGAGAGAGAGAGAG 1451  
 434 IulysProPheAsnIleSerValIlyValLeuProIlyProLeuAsnAla 450





1024 GATAGATATATGATTAATAGACAAATGTAAGGTCAGTGACACG 1073  
 317 SASNAAGGlyGluMetCysAspArgPheGlyCysLeuCysSerProG 334  
 1074 TACCAATCAAGAGAATGTCATCGCTCCCAAGGATGCTCTGCTCTCAAG 1123  
 334 lyTrpGlnGlyLeuGlnCysGluArgGluGlyIleProArgMetThrPro 350  
 1124 CATGCAACAGCGCTGCACCTCTGACAAAGCAAGCCAGCCCAAGCATACACCA 1173  
 351 lysIleValAspLeuProAspHisIleGluValAsnSerGlyLysPheAs 367  
 1174 CAGATAGAGGATTTGATGATCAATTCGATTAACGATGAAATTTTAA 1223  
 367 nProIleCysLysAlaSerGlyTrpProLeuPheThrAsnGluGluMet 384  
 1224 CCCATCTGCAAGCCCTCTGGTGGTCCCTACCTACCTACCTACCTACCT 1273  
 384 hrLeuValLysProAspGlyThrValLeuHisProLysAspPheAsnHis 400  
 1274 CCTAGTGAAGCCAGATGGGACAGTGGTCCCAACCAAAATGACTTCAACTAT 1323  
 401 ThrAspHisPheSerValAlaIlePheThrIleHisArgIleLeuProPr 417  
 1324 ACAGATGCTTCTCAATGATATATTACTTAAACAGCTTACTACTCC 1373  
 417 oAspSerGlyValTrpValCysSerValAsnThrValAlaGlyMetValG 434  
 1374 TCATCTCAGACGCTCTGGCTCTGACGTGTGAACACACAGTGGCTGGGTGG 1423  
 434 luLysProPheAsnIleSerValLysValLeuProLysProLeuAsnAla 450  
 1424 AAAACCCCTTCAACATTCCTGCAAGTCTCCAGAGCCCTCCACGCC 1473  
 451 ProAsnValIleAspThrGlyHisAsnPheAlaValIleAsnIleSerSe 467  
 1474 CCAATGTTGATTGAACATGGAATAAATTGCTATCATCATCATCATCATC 1523  
 467 rGluProIlePheGly 472  
 1524 TGAGCTTACTTGGG 1539

seq\_name: /SIDS1/qcdata/geneseq/emb1/NA2000.DAT:AAZ40638

seq\_documentation\_block:

ID: AAZ40638 standard; cDNA; 4175 BP.

XX AC AAZ40638:

XX DT 07-MAP-2000 (first entry)

XX DE Receptor tyrosine kinase protein encoding cDNA.

XX KW Receptor tyrosine kinase protein; angiogenesis; cardiogen-sis; modu-

XX KW tumorigenesis; ss.

XX OS Mus musculus.

XX PN US5998187-A.

XX PD 07-DEC-1999.

XX PF 23-APR-1997; 9705-0838957.

XX PR 20-JUL-1994; 94US-0278089.

XX PR 30-JUL-1992; 92US-0921745.

XX PR 29-APR-1994; 94US-0235408.

XX PA (MOUN ) MOUNT SINAI HOSPITAL CORP.

XX PI Yamaguchi TP, Breitman J, Dumont DJ, Rossant J, Breitman ML;

XX WPI; 2000-052545/04.

DR

DR P-PSDB: AAY59046.

XX Note: Receptor tyrosine kinase protein involved in angiogenesis, cardiogenesis and tumorigenesis -

XX Disclosure: Fig 1; 123pp; English.

XX The invention provides novel receptor tyrosine kinase protein sequences (AAY59046-48) and nucleic acids encoding the polypeptides. The isolated CC and purified polypeptides are useful for studying the developmental CC expression of a receptor tyrosine kinase protein which may have a role in angiogenesis, cardiogenesis and tumorigenesis. The present sequence CC represents a cDNA encoding a mouse receptor tyrosine kinase protein.

XX Sequence 4175 BP; 1175 A; 942 C; 1047 G; 1011 T; 0 Other;

alignment\_scores:

Quality: 2407.00 Length: 472

Ratio: 5.233 Gaps: 0

Percent Similarity: 97.458 Percent Identity: 91.949

alignment\_block:

US-09-733-764-2\_COPY\_1\_472 x AAZ40638

Align seq 1/1 to: AAZ40638 from: 1 to: 4175

1 MetAspSerLeuAlaSerLeuValLeuCysGlyValSerLeuLeuLeuSe 17  
 124 ATGAGATCTTAAAGGAGCTTAACTTCTTCTGAGATAGCTGCTCTTTA 173  
 17 rGlyThrValGluGlyAlaMetAspLeuIleLeuLeuAsnSerLeuProL 34  
 174 TGGAGTAGTAAAGGCGCCCATGACCTGATCTTGTGATCAATTCCTACCTC 223  
 34 euValSerAspAlaGluThrSerLeuThrCysIleAlaSerGlyTrpAq 50  
 224 TTGTGCTGATGAGGAAATGCTTCTGCTGATCTTCTGCTGCTGCTGCTG 273  
 51 ProHisGluProIleThrIleGlyArgAspPheGluAlaLeuMetAsnGl 67  
 274 CCCCATGAGGCGCATCAATCAATAGAAAGAAATTTGAAAGCTTAAATGAACA 323  
 67 nHisGlnAspProLeuGluValThrGlnAspValThrArgGluTrpAlaL 84  
 324 GCACCAAGATTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTC 373  
 84 yLysValValTrpLysArgGluLysAlaSerLysIleAsnGlyAlaLys 100  
 374 AAAAGTTGTTTGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAG 423  
 101 PheCysGluGlyArgValArgGlyGluAlaIleArgIleArgThrMetIly 117  
 424 TTCTGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 473  
 117 sMetArgGlnGlnAlaSerPheLeuProAlaThrLeuThrMetThrValA 134  
 474 GATCGGCAACAGCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 523  
 134 sPlyGlyAspAsnValAsnIleSerPheLysValLeuIleLysGlu 150  
 524 ACAGGAGAAATATATGTCAATATATCTTTTAAAAAGTGTATTAATTAAGAA 573  
 151 GluAspAlaValIleIleIleIleIleIleIleIleIleIleIleIleIleIle 167  
 574 GAAAGTCACTGATTTACAAAAATGGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 623  
 167 qHisGluValProAspIleLeuGluValHisLeuProHisAlaGlnProG 184  
 624 GCATGAAGTACCTGATATTTTACAAGTTCACTTGGCGCATGCTCAGCCCC 673  
 184 InAspAlaGlyValTyrSerAlaArgTyrIleGlyLysLeuLeuPheThr 200  
 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT





SO Sequence 4364 bp: 1267 A: 965 G: 1056 C: 1099 T: 2 other.

# alignment\_scores:

Quality: 2407 00 Length: 472  
Ratio: 5.243 Gaps: 0  
Percent Similarity: 97.458 Percent Identity: 91.949

## alignment\_block:

us-09-733-764-2\_copy\_1\_472 x AA063730

Align seq 1/1 to: AA063730 from: 1 to: 4364

```

1 MetAspSerLeuAlaSerLeuValLeuCysGlyValSerLeuLeuLeuSe 17
|||||
16 ATGAGCTCTTACCGGGATTATGTCGTGGAGTCACCTTGCCTCTTAA 65
|||||
17 rGlyThrValGluGlyAlaMetAspLeuLeuLeuLeuLeuLeuLeu 34
|||||
66 TGCAGTACAGAGGTGGCAAGCACTGCACTGCACTGCACTGCACTG 115
|||||
34 euValSerAspAlaGluThrSerLeuThrCysIleAlaSerGlyTrpArg 50
|||||
116 TTGGTGTGATGAGTAAATATGCTGACCTGCTGCTGCTGCTGCTG 165
|||||
51 ProHisGluProIleThrIleGlyArgAspPheGluAlaLeuMetAsnGI 67
|||||
166 CGCCATGAGCCCATACATATAGAGAGATTTGAAAGCTTAATGAAACA 215
|||||
67 nHisGlnAspProLeuGluValThrGlnAspValThrArgGluTrpAlaL 84
|||||
216 GCACCAAGATCCACTGGAGGTACTCAAGATGTACCAGAAATGGGCGA 265
|||||
84 yLysValValIrrpLysArgGluLysAlaSerLysIleAsnGlyValatyr 100
|||||
266 AAAAACTCTTGGAGAGACAGAAAGGCCAGTAACATTAATGCTGCTTAT 315
|||||
101 PheCysGluGlyArgValArgGlyGluAlaIleArgIleArgThrMetLy 117
|||||
316 TCTGTGAAAGTGTGATGATGAGATAGATATATAGATATAGATAG 365
|||||
117 sMetArgGlnGlnAlaSerPheLeuProAlaThrLeuThrMetThrValA 134
|||||
366 GATCCCTCAACAAGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 415
|||||
134 sPylsGlyAspAsnValAsnIleSerPheLysLysValLeuIleLysGlu 150
|||||
416 ACAGCGGAGATATATGCTGAACATATCTCTCAAAAAGGCTGTAAATA 465
|||||
151 GluAspAlaValIleThrLysAsnGlySerPheIleHisSerValTrpArg 167
|||||
466 GAAATATGATGATTTTAAATAATGATCTCTCTCTCTCTCTCTCTCT 515
|||||
167 gHisGluValProAspIleLeuGluValHisLeuProHisAlaGlnProG 184
|||||
516 GATCAAGATCTGATGATGATGATGATGATGATGATGATGATGATG 565
|||||
184 IAsnAspIleValIrrpSerAlaArgTrpIleGlyGlyAsnLeuPheThr 200
|||||
566 AGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 615
|||||
201 SerAlaPheThrArgLeuIleValArgArgCysGluAlaGlnIrrpGI 217
|||||
616 TCAATTTTAAATGATGATGATGATGATGATGATGATGATGATGAT 565
|||||
217 gProGluCysAsnHisLeuGlyThrAlaCysMetAsnGlyValCysH 234
|||||
666 GCGCGACTGTAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 715
|||||
234 iSerAspThrGlyGluCysIleCysProProGlyPheMetGlyArgThr 250
|||||
716 ATGAGATATAGCGGGAATGCACTTCCCTCTCTCTCTCTCTCTCTCT 765

```

```

251 CysGluLysAlaCysGluLeuLeuHisThrPheGlyArgTrpCysLysGluArg 267
|||||
766 TGTGATAAATATTATGAGAGAGAGAAATTTTGTATGATGATGATGATG 815
|||||
267 gCysSerGlyGlnGluGlyCysLysSerTrpValPheCysLeuProAspP 284
|||||
816 GTGTACTGGCAATCAGAGAGATGCAAGTCTTATGTGTGTGTGTGTGTGT 865
|||||
284 rGlyGlyCysSerCysAlaThrGlyTrpGlyLysGlyLeuGluCysAsnGlu 400
|||||
866 CTATAGAGGT 915
|||||
301 AlaCysHisProClyPheThrClyProAspCysCysLysCysArgCysSerC 317
|||||
916 GTATGTTCATTTGATTTATATGAGAGAGATGTAAGATGATGATGATG 965
|||||
317 sAsnAsnGlyGluMetCysAspArgPheGlnGlyCysLeuCysSerProG 334
|||||
966 TACCATATGCAAGATATGATGATGATGATGATGATGATGATGATGATG 1015
|||||
334 lYrTrpGlnGlyLeuGlnCysGluArgGluGlyIleProArgMetThrPro 350
|||||
1016 GATGATAGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1065
|||||
351 LysIleValAspLeuProAspHisIleGluValAsnSerGlyLysPheAs 367
|||||
1066 CATATAGAGAGATTTGATGATGATGATGATGATGATGATGATGATGAT 1115
|||||
367 nProIleCysLysAlaSerGlyTrpProLeuProThrAsnGluGluMet 384
|||||
1116 CCGCATATGCAAGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1165
|||||
384 hrLeuValLysProAspGlyThrValLeuHisProLysAspPheAsnHis 400
|||||
1166 CCTATGTAAGGTATATGAGAGATGATGATGATGATGATGATGATGAT 1215
|||||
401 ThrAspHisPheSerValAlaIlePheThrIleHisArgIleLeuProPr 417
|||||
1216 ACAGATGCTGTTCTGAGAGATATATATGATGATGATGATGATGATGAT 1265
|||||
417 oAspSerGlyValIrrpValCysSerValAsnThrValAlaGlyMetValG 434
|||||
1266 TCATTCAGCAGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1315
|||||
434 lLysProPheAsnIleSerValLysValLeuProLysProLeuAsnAla 450
|||||
1316 AAAAGCTTTTAAATTTTGTCTCAAGTCTCTGAGAGAGAGAGAGAGAG 1465
|||||
451 ProAsnValIleCysPheThrGlyHisAsnPheAlaValIleAsnIleSer 467
|||||
1366 CCAATATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1415
|||||
467 rGluProTrpPheGly 472
|||||
1416 TCAGCCTTACTTGTGG 1431

```

seq\_name: /SIS1/qcdata/geneseq/geneseq-emb1/NA1995.DAT:AA091999

seq\_documentation\_block:

ID AA091999 standard; DNA: 4640 bp.

XX AA091999;

XX 21-JAN-1996 (first entry)

XX Mouse tie 2 receptor tyrosine kinase DNA.

XX tie-2; receptor tyrosine kinase; DNA primer; cancer; angiogenesis;

KW vasculogenesis; SS.

XX Mus musculus.

XX Key

XX Location/Qualifiers

FT CDS 341...3712  
 FT /\*tag= a  
 XX W09513387-A1.  
 XX 18-MAY-1995.  
 XX 12-NOV-1994: 94W0-EP03767.  
 XX 12-NOV-1993: 930S-0152552.  
 XX (PLAC ) MAX PLANCK GRS FÖRDERUNG WISSENSCHAFTEN.  
 XX Risau W;  
 XX WPI: 1995-194105/25.  
 DR P-PSDB, AAR73951.  
 XX New tie-2 receptor tyrosine kinase and related nucleic acid - and  
 methods for detecting tie-2 modulators for treating eg cancer,  
 PT associated with angiogenesis and vasculogenesis  
 XX  
 PS Claim 3, Page 42, 81pp, English.  
 CC this DNA may be expressed recombinantly in a host cell. The DNA was  
 CC isolated from mouse brain and can be used as a probe to detect  
 CC related genes of other animals or expressed in host cells e.g E  
 CC coli. The expressed protein may be used to treat diseases, or  
 CC processes, associated with angiogenesis and vasculogenesis, or  
 CC cancer. Cells that express the protein are used in screening  
 CC procedures and recombinant protein can be used for affinity  
 CC purification of tie-2 ligand. The DNA may be used diagnostically to  
 CC detect gene expression.  
 XX  
 SQ Sequence 4640 BP: 1403 A: 1041 C: 1147 G: 1149 T: 0 other:  
 alignment\_scores:  
 Quality: 2407.00 Length: 472  
 Ratio: 5.233 Gaps: 0  
 Percent Similarity: 97.458 Percent Identity: 91.949  
 alignment\_block:  
 US-09-733-764-2\_copy\_1\_472 x AAG91999 ..  
 Align seq 1/1 to: AAG91999 from: 1 to: 4640  
 1 MetAspSerLeuAlaSerLeuValIleCysGlyValSerLeuLeuLeu 17  
 341 ATGGACTCTTTAGCGCGCTTACTTCTCTGCGAGTCAGCTTCCTCTTTA 390  
 17 rGlyThrValGluGlyAlaMetAspLeuIleLeuIleAsnSerLeuProl 34  
 391 TGGAGTAGTAGAGGCGGACCGACCGACCTTGATCAATTCCTCACCCTC 440  
 34 cuValSerAspAlaGluThrSerLeuThrCysTleAlaSerGlyTyrArg 50  
 441 TTGTGTCTATGAGGAGAAATATCTCTATCTATCTATCTATCTATCT 490  
 51 ProHisGluProlIleThrIleClyArgAspPheGluAlaIleuMetAsnCl 67  
 491 GCGCATGAGCGGACATACCATAGGAGGAGGACTTTGAGGCGCTTAATGAGCA 540  
 67 mHisGlnAspProLeuGluValThrGlnAspValThrArgGluTrpAlaL 84  
 541 GATCTAAGATCTATCTAGCTTACTCAAGATGAGGAGGAGGAGGAGGAG 590  
 84 yLysValValTrpLysArgGluLysAlaSerIlystIleAsnGlyAlaTyr 100  
 591 AAAAAGTTCTTGGAAAGAGAGAAAGGCGGCTAAGATTATATGCTGCTAT 640  
 101 PheCysGluGlyArgValArgGlyGluAlaIleCargIleArgThrMetIly 117  
 1172 TTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 1590

641 TTCCTGAGGCTGGAGTTCAGGAGACAGGCTATAGGATACAGGACCA 690  
 117 sMetArgGlnGlnAlaSerPheLeuProAlaThrLeuThrMetThrValA 134  
 691 GATGCTCTAAATAGGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 740  
 134 sPheGlyAspAsnValAsnIleSerPheLysLysValLeuIleLysGlu 150  
 741 ACAGGAGAGATATGTCAGATATCTTCAAAAGGCTGTAATTAAGAGAA 790  
 151 GluAspAlaValIleTyrLysAsnGlySerPheIleIleHisSerValProAr 167  
 791 CAAGTCAGTGATTTACAAAATGGCTCTCTCATCCACTCAGTGGCCCG 840  
 167 qHisGluValProAspIleLeuGluValHisLeuProHisAlaGlnProG 184  
 841 GCAICAAATATCTATATTTTAAAGCTCAATTGCGGATAGCTATAGAGCC 890  
 184 LAspAlaGlyValTyrSerAlaGlyTyrIleGlyGlyAsnGluLeuHe 200  
 891 AATAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 940  
 201 SerAlaPheThrArgLeuIleValArgCysGluAlaGlnLysTrpG1 217  
 941 TCAGCTTCAGCAGGCTGCAATGTCGAGAGAGAGAGAGAGAGAGAGAGAG 990  
 217 YProGlyCysAsnHisLeuCysThrAlaCysMetAsnAsnGlyValCysH 234  
 991 GTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1040  
 234 LAspAspIleGlyCysIleCysIleCysIleCysIleCysIleCysIle 250  
 1041 ATGAGATATAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1090  
 251 CysGluLysAlaCysGluLeuHisThrPheGlyArgThrCysLysGluAr 267  
 1091 TGTGAGAAAGCTGTCAGGCTGATACATATGCGAGAGAGAGAGAGAGAGAG 1140  
 267 qCysSerIleGlnIleGlyLysLysSerTyrValPheCysLeuProAspP 294  
 1141 GTTAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1190  
 284 rGlyCysSerCysAlaThrGlyTrpLysGlyLeuGlnCysAsnGlu 300  
 1191 CTAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1240  
 301 AlaCysHisProGlyPheTyrGlyProAspCysLysLeuArgCysSerCy 317  
 1241 GCATGCGCTATCTGCTTACTAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1290  
 317 sAsnAsnGlyGluMetCysAspArgPheGlnGlyCysLeuCysSerProG 334  
 1291 TATCAATGAGAGATATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1340  
 334 LyrTrpGlnGlyLeuGlnCysGluArgGlyIleProArgMetThrPro 350  
 1341 GATGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1390  
 351 LysIleValAspLeuProAspHisIleGluValAsnSerGlyLysPheAs 367  
 1391 CAGATAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440  
 367 nProIleCysLysAlaSerGlyTrpProLeuProThrAsnGluGluMet 384  
 1441 CCGATATCTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1490  
 384 hLeuValIleProAspGlyThrValIleuHisProLysAspPheAsnHis 400  
 1491 CCGTAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1540  
 401 ThrAspHisPheSerValAlaIlePheThrIleHisArgIleLeuProPr 417  
 1541 ACAGATGCTTCTGAGTGGGCAATATTCAGTCAAGGAGGAGGAGGAGGAGGAG 1590





```

88 rplysArgGlyCysAlaSerLysLysAsnGlyAlaThrPheCysGluGly 104
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
316TCCAGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAGAT 357
105 ArgValArgGlyGluAlaIleArgIleArgThrMetLysMetArgGlnC 121
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
358 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 407
121 nAlaSerPheLeuProValThrLeuThrMetThrValAspLysGlyAspA 138
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
408 AGCCAGCAGCTTCCACAGAGGTCACACACACTGTGACAAAGAGGTGACA 457
138 snValAsnLysSerPheLysValIleLysGlnGluAspAlaVal 154
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
458 CCGCTGTATTTCTGACGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 507
155 IlePylLysAsnGlySerPheLysSerValProArgHisGluValPro 171
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
508 TCCAGAGAGTACAGAGTACAGAGTACAGAGTACAGAGTACAGAGTAC 557
171 oAsp.....IleLeuGluValIleLysLeuProLysAlaGlnProGlnA 185
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
558 GCATGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 601
185 spAlaGlyValTyrSerAlaArgGlyTyrIleGlyGlyAsnLeuPheThrSer 201
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
602 CGAGGAGATATTAATGAGATATTAATGAGATATTAATGAGATATTAAT 651
202 AlaPheThrArgLeuLeuValArgArgGlyGluAlaGlnLysTyrGlyPyr 218
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
652 CCGCTTTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 701
218 GlnLysAsnLysLeuGlyThrAlaCysMetAsnAsnLysValCysHisG 235
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
702 AGGCTCTACCAAGGAGTACAGGAGTACAGGAGTACAGGAGTACAGGAGT 751
235 LuAspThrGlyGluCysLysCysProProGlyPheMetGlyArgThrCys 251
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
752 ACATGAGGAGGAGTATGATGATGATGATGATGATGATGATGATGATGAT 801
252 GlutLysAlaLysGlnLeuLysThrPheGlyArgThrCysGluArgCys 268
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
802 CAACAGGAGTACAGGAGTACAGGAGTACAGGAGTACAGGAGTACAGGAGT 851
268 sSerGlyGluGlyCysLysSerTyrValPheCysLeuProAspProT 285
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
852 GGCAGGATATACAGGAGTACAGGAGTACAGGAGTACAGGAGTACAGGAGT 901
285 rGlyCysSerCysAlaThrGlyTyrPheGlyPheGlnGlySerAsnGluAla 301
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
902 ATGAGTGTATTGATGATGATGATGATGATGATGATGATGATGATGATGAT 951
302 CysHisProPhePheGlyProAspGlyLysLysLeuArgCysSerCysAs 318
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
952 TGTGCCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1001
318 nAsnGlyGluMetCysAspPheGlnGlyGlyCysLeuCysSerProGlyT 335
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1002 GAATGCTGGTACTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1051
335 rGlnGlyGluGlnCysGluArgIleGlyIleProArgMetThrPheLys 351
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1052 GGCATGAGGAGTACGTGAGTACGTGAGTACGTGAGTACGTGAGTACGTGAG 1092
352 IleValAspLeuProAspHisIleGluValAsnSerGlyLysPheAsnPro 368
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1093 AGCTCAACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1142
368 oile....CysLysAlaSerGlyTyrProLeuProThrAsnGlnGluMetT 384
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1143 GATCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1192

```

```

384 hrLeuValLysProAspGlyThrValLeuHisProLysAspPheAsnHis 400
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1193 AGTATATTAATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTAT 1242
401 ThrAspHisPheSerValAlaIlePheThrIleHisArgIleLeuProPr 417
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1243 GAGCCAG 1292
417 oAspSerGlyValTyrValCysSerValAsnThrValAlaGlyMetValG 444
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1293 GCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1342
434 LuLysProPheAsnLysSerValLysValCysProLysProLeuAsnAla 450
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1343 GCGGCGGCTTCAAGGTCAGAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 1402
451 ProAsnValIle 454
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1393 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1404

seq_name: /SIDS1/acqdata/geneseq/geneseq-emb1/NA2001B.DAT:AA579371
seq_documentation_block:
ID AA579371 standard: cDNA; 6249 bp.
XX
AC AA579371;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #15175.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PP 31-MAR-2000; 2000US-0540217.
PP 23-AUG-2000; 2000US-0649167.
XX
FA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
WP1, 2001-039362/73.
DR P-PSDB: ABG15184.
XX
New isolated polynucleotide and encoded polypeptides, useful in
diagnostics, forensics, gene mapping, identification of mutations
responsible for genetic disorders or other traits and to assess
biodiversity -
XX
Claim 1. SEQ ID NO 15175; 103pp; English.
XX
The invention relates to isolated polynucleotide (I) and
polymers (II) sequences. (I) is useful as hybridisation probes,
polymerase chain reaction (PCR) primers, oligomers, and for chromosome
and gene mapping, and in recombinant production of (II). The
polynucleotides are also used in diagnostics as expressed sequence tags
for identifying expressed genes. (I) is useful in gene therapy techniques
to restore normal activity of (II) or to treat disease states involving
(II). (II) is useful for generating antibodies against it, detecting or
quantitating a polypeptide in tissue, as molecular weight markers and as
a food supplement. (II) and its binding partners are useful in medical
imaging of sites expressing (II). (I) and (II) are useful for treating
disorders involving aberrant protein expression or biological activity.
The polypeptide and polynucleotide sequences have applications in
diagnostics, forensics, gene mapping, identification of mutations

```



CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. AA564197-AA594564 represent novel human  
 CC diagnostic coding sequences of the invention.  
 CC Note: The sequence data for this patent did not appear in the printed  
 CC specification, but was obtained in electronic format directly from wipo  
 CC at ftp.wipo.int/pub/published\_pat\_sequences.

XX Sequence 6249 BP: 2146 A; 1308 G; 1403 G; 1465 T; 0 other.

#### alignment\_scores:

Quality: 688.50 length: 141  
 Ratio: 5.062 Gaps: 1  
 Percent Similarity: 96.454 Percent Identity: 95.745

#### alignment\_block:

US-09-733-764-2\_COPY\_1\_472 x AA579371

Align seg 1/1 to: AA579371 from: 1 to: 6249

```

69 GlnAspProLeuGluValThrGlnAspValThrArgGluTrpAlaLysLy 85
 ::
52 CAGTATCGCTGTAAGTACTCAAGATGAGACAGACAAAGGCTAAAAA 131
 ::
85 svaValTrpIysArgGluLysAlaSerLysIleAsnGlyAlaTyrPheC 102
 ::
102 ASTGTGTGGAAGAAACAATAAGAGCTATGATGATCAATGGTGTATTTC 151
 ::
102 ysGluGlyArgValArgGlyGluAlaIleArgIleArgThrMetLysMet 118
 ::
152 GTGAAGCGCGAGTTCGAGGAGAGCAATCAGGATACCAACCAATCAAGATG 201
 ::
119 ArgGlnGlnAlaSerPheLeuProAlaThrTrpMetThrValAspLy 135
 ::
202 CCTCAACAA.....GCTACTTTAACTATGACTGTGGACAA 236
 ::
135 sGlyAspAsnValAsnIleSerPheLysValIleLysGluGluA 152
 ::
237 GGGAGATACAGTGGAAATATCTTTCAAAAGGTATTGATTAAAGAAGAG 286
 ::
152 sPaIaValIleTyrIysAsnGlySerPheIleHisSerValProArgHis 168
 ::
287 ATGCAGTCAATTTACAAAAATGTTCTCTTCATCCATCCAGTCCCGCGCAT 336
 ::
169 GluValProAspIleLeuGluValHisLeuProIleHisAlaGlnProGlnAs 185
 ::
337 GAAGTACCTGATATCTAGAAAGTACACTGCTGCTCATGCTCAGGCTCAAGA 386
 ::
185 pAlaGlyValTyrSerAlaAATTTCTTCTTCTTCTTCTTCTTCTTCTTCT 202
 ::
387 TGTGGAATGATATGAGGATATATATATATATATATATATATATATATAT 436
 ::
202 IapheThrArgGluValValArg 209
 ::
437 CCTTCACCAGCTCATAGTCCCG 459

```

seq\_name: /SDS1/seqdata/geneseq/geneseq\_emb1/NA1943 DAT-AAQ47065

#### seq\_documentation\_block

AAQ47065 standard: cDNA; 3713 BP.

XX AAQ47065:

XX AC

XX AT 12-JAN-1994 (first entry)

XX DE tie truncated receptor kinase cDNA.

XX KW Tie; receptor; tyrosine kinase; endothelial cell; immunoglobulin; Ig;  
 KW epidermal growth factor; EGF; repeat; blood vessel; atherosclerosis;  
 KW thromboembolic disease; neoplastic disease; tumour angiogenesis;  
 KW wound healing; inflammatory disease; ss.

#### Homo sapiens.

XX Key Location/Qualifiers  
 XX CDS 37..3321  
 XX /\*tag= a  
 XX sig\_peptide 37..99  
 XX /\*tag= b  
 XX mat\_peptide 100..3318  
 XX /\*tag= c

XX W09314124-A.

XX 22-JUL-1993.

XX 08-JAN-1994; 93WO-FI00006.

XX 09-JAN-1992; 92US-0817800.

XX (UYHE-) UNIV HELSINKI HOLDING LTD.

XX Alitalo K, Armstrong E, Korhonen J, Makela TP;  
 XX Partanen J;

XX WP: 1993 24352/30.

XX P-PSDB; AAR39821.

XX Nucleic acid encoding tyrosine kinase receptor - for regulating  
 PT tie activity in diagnosis and treatment of neoplastic diseases  
 PT involving tumour angiogenesis, wound healing etc.

XX Disclosure: Page 53-56; 73pp; English.

XX The sequence given in AAQ47064 represents the full length 'tie' receptor  
 CC kinase cDNA. The sequence given in AAQ47065 represents a truncated  
 CC version of this, wherein nucleotides corresponding to positions 676-  
 CC 807 of the first sequence are absent. Tie is an endothelial cell  
 CC receptor tyrosine kinase and contains immunoglobulin (Ig-) and  
 CC epidermal growth factor (EGF-) like repeats. The tie gene is  
 CC expressed in the endothelial cells of blood vessels. Tie levels may  
 CC indicate onset/progression of diseases involving endothelial cells  
 CC and their tie receptors, such as neoplastic diseases involving tumour  
 CC angiogenesis, wound healing, thromboembolic diseases, atherosclerosis  
 CC and inflammatory diseases.

XX Sequence 3713 BP: 722 A; 1142 C; 1107 G; 742 T; 0 other;

#### alignment\_scores:

Quality: 675.50 length: 454  
 Ratio: 2.474 Gaps: 3  
 Percent Similarity: 60.132 Percent Identity: 33.040

#### alignment\_block:

US-09-733-764\_2\_COPY\_1\_472 x AAQ47065

Align seg 1/1 to: AAQ47065 from: 1 to: 3713

```

14 LeuLeuLeuSerGlyThrValGlyAlaMetAspLeuIleLeuAs 30
 :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
73 GAGTCTTGGCTTCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 122
 :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
30 nSerLeuProLeuValSerAspAlaGluThrSerLeuThrCysTle.... 45
 :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
123 CAATTTGAGATCAAGAAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 172
 :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
46 AlaSerGlyTrpArgProHisGluPro 54
 :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
173 GGCAGGCG 216
 :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
55 IieThrIleGlyArgAspPheGluAlaLeuMetAsnGlnHisGlnAspPr 71
 :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::
217 GAGTCTTGGCTTCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 266
 :::::::::::::: :::::::::::::: :::::::::::::: ::::::::::::::

```







insotcicoids, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (AB161676-AB163511), expressed DNA sequences (AB161840-AB161875) and the encoded proteins (ABBS7737-ABBT2072).

The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from [www.ncbi.nlm.nih.gov/pubmed/11111111](http://www.ncbi.nlm.nih.gov/pubmed/11111111).

|                     |        |                          |
|---------------------|--------|--------------------------|
| alignment_scores:   |        |                          |
| Quality:            | 331.00 | Length: 353              |
| Ratio:              | 1.891  | Gaps: 14                 |
| Percent Similarity: | 49.575 | Percent Identity: 27.195 |

alignment block.

Align seg 1/1 to reverse of: ABL04034 from: 1 to: 17478

47 SerGlyTyrArgProHisIleuProIleThrIleGly  
|||||  
5115 TCCGATGGAAAGCA TGGCGGACAGTGGCAGAGGAC TGTGCCGTGGCAGA 5066

59 . . . . . ArgAspPheGluAlaIleuMetAsnGlnHisC 69  
|||||  
5065 ACGAGGAAAAATGTTAAATAGAAATGGGAGATGATATGTAATGATAGG 5016

69 InAspProLeuGluValThrGlnAspValThr . . . . . ArgGlu 81  
||| . . . . . |||  
5015 TCGATGGGGGAGAGGAGGAGCAAAAGAGTGAGGAGGAGAGGAGGAGGAG 4966

82 TrpAlaLysLysValValTrpLysArgGluLysAlaSerLysIleAsnGlu 98  
||| . . . . .  
4965 TCGATGCCAAGA . . . . . 4964

98 yAlaTyrPheCysGluGlyArgVal . . . . . ArgGlyGluAlaIleArgIleA 114  
||| . . . . . |||  
4964 GAGCTGGAGTGGCTATAGAGAGAGGCTGATGATGAGAGAGAGAGAG 4917

114 yGThrMetLysMetArgGlnGlnAlaSerPheLeuProAlaThrLeuThr 130  
:. . . . . |||||  
4916 ATCACC GGCCAGTGGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAG 4857

131 MetThrValAspLysGlyAspAsnValAsnIleSerPheLysLysValle 147  
||| . . . . . |||  
4866 GCACCACTG TCACCTCAACACCTACGGATTCACCTCAACCTCAACCTCA 4820

147 uIleLysGlnGlu AspAlaValIleCylLys . . . . . AspGlySerPheIle 162  
||| . . . . . |||||  
4819 CGGATTCGGATGAGATGATGATGATGATGATGATGATGATGATGATG 4770

163 HisSerValProArgHisGluValProAspIleLeuGluValHisLeuPr 179  
4770 . . . . . 4770

179 oHisAlaGlnProGlnAspAlaGlyValTyrSerAlaArgTyrIleGlyC 196  
|||||  
4769 TCGATTCGGATTCGGATTCGGATTCGGATTCGGATTCGGATTCGGATTC 4734

196 yAsnLeuPheThrSerAlaPheThrArgLeuIleValArgArgCysGlu 212  
||| . . . . . |||||  
4733 . . . . . CGGATTCGGATTCGGATTCGGATTCGGATTCGGATTCGGATTC 4722

213 AlaGlnLysTrpGlyProGlyCysAsnHisLeuCysThrAlaCysMet . . 228  
||| . . . . . |||  
4721 CGGAACAATAAGGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 4672

229 . AsnAsnGlyValCysHisGlnAspThrGlyLysLysLysProProG 245  
||| . . . . . |||||  
4671 GCACCAAGGATTCGTCACCAAGGATTCGTCACCAAGGATTCGTCACCAAG 4622





34 euValSerAspAlaGluThrSerLeuThrCysIleAlaSerGlyTrpArg 50  
|||||  
249 TTCTATCTGATGCTGGAACATCTCTCACTGCAATCCCTCTGGGTGGGC 298  
51 ProHisGluProGluThrIleGlyArgAspPheGluAlaLeuMetAsnG1 67  
|||||  
299 GAGCATGAGGCGATGATGATAGAGAGAGACATTTGAAGGCTTAATGAACCA 348  
67 nHisGluAspProGluGluValThrGlnAspValThrArgGluTrpAlaIat 84  
|||||  
349 GCACGAGGATGCTGCAACTTACTCAAGATGTCACGACAGAAATGGCGCTA 398  
84 yLysValValTrpLysArgGlnLysAlaSerLysIleAsnGlyAlaTrp 100  
|||||  
399 AAAAAGTTCTTTGGAGAGAGAGAGAAAAGCTAGTAAAGATCAATGGCTTAT 448  
101 PheCysGluGluValArgGlyGluAlaIleCargIleArgThrMetLys 117  
|||||  
449 TTCTGTGAGAGATGAGTGTGAGAGAGAGATGAGATGAGATGAGATGAG 498  
117 sMetArgGluThrAlaSerPheLeuProAlaThrLeuThrMetThrValA 134  
|||||  
499 CATCCCTCAACAAGCTTCTCTCCACCACTACTTTAACTATGACTGTGG 548  
134 sPlyGlyAspAsnValAsnIleSerPheLysValLeuIleLysGlu 150  
|||||  
549 ACAAGGAGATTAACGTGAAATATCTTCAAAAAGCTATTCATTAAGAAA 598  
151 GluAspAlaValIleThrLysAsnGlySerPheIleHisSerValProAr 167  
|||||  
599 GAAGATGCTACTATTATTAATAATGCTTCTTATATATATATATATATAT 648  
167 qHisGluValProAspIleGluGluValHisLeuProHisAlaGlnProG 184  
|||||  
649 GCATGAAGTACCTGATATCTGAGGAGATATAGAGGAAACCTCTCTACC 698  
184 InAspAlaGlyValTyrSerAlaLeuThrIleGlyGluValPheThr 200  
|||||  
699 AGATGCTGAGTACCTGAGGAGATATAGAGGAAACCTCTCTACC 748  
201 SerAlaPheThrArgLeuIleValArgAspCysGlnAlaGlySerTrpG1 217  
|||||  
749 TGGGCTTCAACAGATATATATATATATATATATATATATATATATATAT 798  
217 yProGluCysAsnHisLeuGlySerAlaCysMetAsnAsnGlyValCysH 234  
|||||  
799 ACCTGAATGCAACCATCTCTGAGGAGATATAGAGGAAACCTCTCTACC 848  
234 IsGluAspThrGlyGluGlySerPheProGlyPheMetGlyValThr 250  
|||||  
849 ATGAAGATGCTGAGGAGATATAGAGGAAACCTCTCTACC 898  
251 CysGluValAlaCysGlnAlaHisThrPheGlyArgThrCysLysGluAr 267  
|||||  
899 TGTGATTAAGGCTTGTGAAATATATATATATATATATATATATATATAT 948  
267 qCysSerGlyGluGluCysLysSerTrpValPheCysLeuProAspP 284  
|||||  
949 GTGCACTGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 998  
284 reTyroLysSerPheValThrGlyTrpLysIleLeuGlnCysAsnGlu 300  
|||||  
999 CCTATGGGTCTCTCTGTCACAGGCTGCAAGGCTCTGCACTGCAATGAA 1048  
301 AlaCysHisProGlyPheThrGlyProAspPheLysLysLysLysLysLys 317  
|||||  
1049 GCAATGATATATATATATATATATATATATATATATATATATATATAT 1098  
317 sAsnAsnGlyGluMetCysAspArgPheGlnGlyCysLeuCysSerProG 334  
|||||  
1099 CAACAATGGGAGATGCTGAGATCTTCCACAGATGCTCTGCTCTCCAG 1148  
334 IyTrpGlnGlyLeuGlnCysGluArgGluGlyIleProArgMetThrPro 350

1149 GATGGAGGAGTCTAGTGTATAGATATAGAGATATATGAGGATGACCCA 1198  
351 LysIleValAspLeuProAspHisIleGluValAsnSerGlyLysPheAs 367  
|||||  
1199 AGATAGTGTATTTGATGATATATAGAAATTAATATGCTTAATTTAA 1248  
367 nProLysCysLysAlaSerGlyTrpProLeuProThrAsnGluGluMet 384  
|||||  
1249 TCCCATTTGCAAAATCTGCTGGTGGGCTACCTACTAATGCAAGAAATCA 1298  
384 hrLeuValLysProAspGlyThrValLeuHisProLysAspPheAsnHis 400  
|||||  
1299 GCTGTGTAAGAGGATGATGATATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1348  
401 ThrAspHisPheSerValAlaIlePheThrIleHisArgIleLeuProP 417  
|||||  
1349 AGGATGATTTTATATATATATATATATATATATATATATATATATAT 1398  
417 oAspSerGlyValTrpValCysSerValAsnThrValAlaGlyMeIValG 434  
|||||  
1399 TCAATATAGAGATCTGGTGTGTAGTGTGAACACATCTGGTGGGATGGTGG 1448  
434 LuLysProPheAsnIleSerValLysValLeuProLysProLeuAsnAla 450  
|||||  
1449 AAAAGGCTTCAACATTTCTGTAAAGTCTTCAAAAGGCTGTAATGCC 1498  
451 ProAsnValIleAspThrGlyHisAsnPheAlaValIleAsnIleSerSe 467  
|||||  
1499 TTAATAATGATTTATATATATATATATATATATATATATATATATATAT 1548  
467 rGluProTyrPheGly 472  
|||||  
1549 TGAGGCTTACTTTGGG 1564  
seq\_name: /cqn\_2\_5/ptodata/1/usa/PC/US\_COMB.seq:PC1-US93-06093-1  
seq\_documentation\_block:  
; Sequence 1, Application PC/US9306093  
; GENERAL INFORMATION:  
; APPLICANT: Ziegler, Steven F.  
; TITLE OF INVENTION: NOVEL TYROSINE KINASE  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Immunex Corporation  
; STREET: 51 University Street  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: US  
; ZIP: 98101  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.05  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PC1/US93/06093  
; FILING DATE: 19930625  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; AFFILIATION NUMBER: US 07,905,600  
; FILING DATE: 26-JUN-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Soose, Kathryn A.  
; REGISTRATION NUMBER: 32,172  
; REFERENCE/DOCKET NUMBER: 2609  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206) 587-0430  
; TELEFAX: (206) 233-0644  
; TELEX: 756822  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 4138 base pairs



```

? TYPE: NUCLEIC ACID
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULAR TYPE: cDNA to mRNA
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 149..3523
PCT-US93-06093-1

```

```

alignment_scores:
 Quality: 2695.00 Length: 472
 Ratio: 5.519 Gaps: 0
 Percent Similarity 100.000 Percent Identity 100.000

```

```
alignment_block:
```

```
US-09-733-764-2_COPY_1_472 x PCT-US93-06093-1
```

```
Align seq 1/1 to: PCT-US93-06093-1 from: 1 to: 4138
```

```

1 MetAspSerLeuAlaSerLeuValLeuCysGlyValSerLeuLeuLeuSe 17
|||||
149 ATGCACCTCTTAGCCAGCTTAGTCTCTGTCGAGTCAGCTGCTCTCTTC 198
|||||
17 rGlyThrValGluGluAlaMetAspLeuLeuLeuLeuLeuLeuProL 34
|||||
194 TGGAACTGTGCGAAGCTGCCATGGACTTGATCTTGATCAATTCCTTACC 248
|||||
34 cuValSerAspAlaGluThrSerLeuThrCysIleAlaSerGlyTrpAr 50
|||||
249 TTATATCTGATGCTTAAATATTTTAACTGATTTGCTGCTGCTGCTG 298
|||||
51 proHisGluProIleThrIleGlyArgAspPheGluAlaIleuMetAsn 67
|||||
299 CCCCATGAGCCCATCACCATAGTAAGGAGCTTTGAAGCCTTAATGAAG 348
|||||
67 nHisGluAspProLeuGluValThrGlnAspValThrArgGluTrpAla 84
|||||
349 GCACCTAGATCTCCCTGGAGTAACTCAAGATGTCAGACAGACATGGC 398
|||||
84 yHisValValThrPheArgGluAlaSerIleAsnGlyAlaTrp 100
|||||
399 AAAAAGTTGCTTTGCAACAGAAAGAGTTATGTAAGATCAATGGTGT 448
|||||
101 PheCysGluGlyArgValArgGlyGluAlaIleArgIleArgThrMet 117
|||||
449 TTCCTGCAAGGGCGAGTTCGAGGAGAGGCAATCAGATACGAACCATGA 498
|||||
117 sMetArgGlnGluAlaSerPheLeuProAlaThrLeuThrMetThrVal 134
|||||
499 GATCGGTCACCAAGCTTCCTTCCCTACCAGCTACTTTAACTATGACTGT 548
|||||
134 sPheGlyAspAsnValAsnIleSerPheIleCysValLeuIleCysGlu 150
|||||
549 ACAAGGAGATAAATGTAATATATCTTTCAAAAAGGTATGATTAAAGAA 598
|||||
151 GluAspAlaValIleThrIleCysAsnGlySerPheIleHisSerVal 167
|||||
599 GAACATCCAGTGTATACAAAAATGGTTCCTTCATCCATTCAGTCCCCG 648
|||||
167 qHisGluValProAspIleLeuGluValHisLeuProHisAlaGlnPro 184
|||||
649 GCAAGAAGTACCTGCAATATATTAAGATACCTGCTGCTGCTGCTGCT 698
|||||
184 hAspAlaGlyValThrSerAlaArgThrIleCysIleCysLeuLeuThr 200
|||||
699 AGAATGTGAGTGTATATGAGTATATATATATATATATATATATATAT 748
|||||
201 SerAlaPheThrArgGluLeuValArgGlyCysGluAlaGlnIleTrp 217
|||||
749 TCGGCTTCACCAAGGATGATAGTCCGAGACAGTGAAGTCCCAAGAGG 798
|||||
217 yProGluCysAsnHisLeuCysThrAlaCysMetAsnGlyValCysH 234

```

```

|||||
799 ACCTGAATGCAACCATCTCTGTAAGTGTATGAACAATGGTGTCTGGC 848
|||||
234 iScIAspPheThrGlyCysIleCysProProGlyPheMetClyArgThr 259
|||||
849 ATGAAGATATGAGAAATGATTTTGGCTGCTGAGGTTATGAGGAAAG 898
|||||
251 CysGluLysAlaCysLeuLeuHisThrPheGlyArgThrCysLysGlu 267
|||||
899 GTGCAGAAAGGCTTGTCACTGCACAGCTTGGCAGAACCTGTAAGAAAG 948
|||||
267 gCysSerGlyGluGluCysLysSerTyrValPheCysLeuProAspF 284
|||||
949 GTGATGAGTGAAGAAAGTATGAAAGTATGATGATGATGATGATGAT 998
|||||
284 rGlyThrCysSerCysAlaThrGlyTrpLysCysLeuGlnCysAsnGlu 309
|||||
999 GATATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1048
|||||
301 AlaCysHisProGlyPheThrGlyProAspCysLysLeuArgCysSer 317
|||||
1049 GCATGCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1098
|||||
317 sAsnAsnGlyGluMetCysAspArgPheGlnGlyCysLeuCysSerPro 334
|||||
1099 CAAATATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1148
|||||
334 lYTrpGlnGlyLeuGlnCysGluArgGlyCysIleProArgMetThr 359
|||||
1149 GATGATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1198
|||||
351 LysIleValAspLeuProAspHisIleGluValAsnSerGlyLysPhe 367
|||||
1199 AAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1248
|||||
367 nProIleCysLysAlaSerGlyTrpProLeuProThrAsnGluGlu 384
|||||
1249 TCCATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1298
|||||
384 hLeuValLysProAspGlyThrValLeuHisProLysAspPheAsnHis 400
|||||
1299 CCTGCTGAAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1348
|||||
401 ThrAspHisPheSerValAlaIlePheThrIleHisArgIleLeuPro 417
|||||
1349 ACCGATCATCTTCAGTAGGCTATATTCACCAATCCACCGGATCCCT 1398
|||||
417 oAspSerGlyValTrpValCysSerValAsnThrValAlaGlyMetVal 434
|||||
1399 TGAATGAGGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1448
|||||
434 lLysProPheAsnIleSerValLysValLeuProLysProLeuAsnAla 450
|||||
1449 AAAAGGCTTCAATATTTGTTAAATCTCTCAAAAGGCTTGAATGCT 1498
|||||
451 ProAsnValIleAspThrGlyHisAsnPheAlaValIleAsnIleSer 467
|||||
1499 CCAACCTGATTCACACTGCACATACTTTGCTGCTGCTGCTGCTGCT 1548
|||||
467 rGluProTyrPheGly 472
|||||
1549 TGACCTTATTTTGGG 1564

```

```
seq_name: /cgn2_6/ptdata/1/lna/5A_COMH seq-US-07-934-393B-1
```

```
seq_documentation_block:
```

```
; Sequence 1. Application US/07934393B
```

```
; Patent No. 5466596
```

```
; GENERAL INFORMATION:
```

```
; APPLICANT: BREITMAN, MARTIN L.
```

```
; APPLICANT: DOMONT, DANIEL
```

```
; APPLICANT: GRADWIL, GERALD G.
```

```
; TITLE OF INVENTION: TISSUE SPECIFIC TRANSCRIPTIONAL
```





```

167 qhlsGluValProAspIleLeuGluValHisLeuProHisAlaGlnProG 184
|||||
624 GCAIGAAAGTACCAATCAATTTACAAAGTTTACTGGCCCAAGTCAGCC 673
|||||
184 InAspAlaGlyValTyrSerAlaAspTyrIleGlyGlyAsnLeuPheThr 200
|||||
674 AGGATGCTGGTGGTCAATCGAGTCAAGTAAATAGAGAGAAATCTGTTC 723
|||||
201 SerAlaPheThrArgLeuIleValArgTyrCysGluAlaGlnTyrPhe 217
|||||
724 TCAGCTTCAGCAAGGTGATTTTGGCAATGTGAAGTTCAGAACTGGAG 773
|||||
217 yProGluCysAsnHisLeuCysThrAlaCysMetAsnAsnGlyValCysH 234
|||||
774 GCGGACGTGAGCCGCTGGTACTATCTGCAAGCAACAATGGAGTCTGCC 823
|||||
234 isGluAspThrIleGlyCysAlaCysProProGlyPheMetGlyArgThr 250
|||||
824 ATGAAATATATGAGGAATATATGGATCTCTGGCTTATATGAGAAATGA 873
|||||
251 CysGluLysAlaCysGluLeuHisThrPheGlyArgThrCysLysGluAr 267
|||||
874 TGTGAGAAAGATTCTGAGGAGGAAATATTGGAGAGAAATGAGAAAGAG 923
|||||
267 GCGSAspGlyGluGluGlyCysLysSerTyrValPheCysLeuProAspP 284
|||||
924 GTCTAGTGGACCAAGCAAGCAAGCTTATGCTGCTGCTGCCAGACC 973
|||||
284 rGlyGlyCysSerCysAlaTyrGlyTyrLysGlyLeuGluCysAsnHis 300
|||||
974 GTTACGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1023
|||||
301 AlaCysHisProGlyPheCysGlyProAspCysLysLeuAlaCysSerCy 317
|||||
1024 GATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1073
|||||
317 sAsnAsnGlyGluMetCysAspArgPheGlnGlyCysLeuCysSerProG 334
|||||
1074 TACCAATCAAGACATAATGACATGCTGCTGCTGCTGCTGCTGCTGCTG 1123
|||||
334 lyTrpGlnGlyLeuGlnCysGlyArgGlnGlyIleProArgMetThrPro 350
|||||
1124 GATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1173
|||||
351 lysIleValAspLeuProAspHisIleGluValAsnSerGlyLysPheAs 367
|||||
1174 CAGATAGAGATTTCGACATCAATTAAGTAAATAGTGAATAATTTAA 1223
|||||
367 nProIleCysLysAlaSerGlyTrpProLeuProThrAsnGluGluMet 384
|||||
1224 GCGCATCTGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1273
|||||
384 hrLeuValLysProAspGlyThrValLeuHisProLysAspPheAsnHis 400
|||||
1274 GCGTAGTGAAGATTAATGAGAAATGACAAATGAAATTAATTAATTA 1323
|||||
401 ThrAspHisPheSerValAlaIlePheThrIleHisArgIleLeuProPr 417
|||||
1324 ACAGATGGTTTCTCAGTGGACAAATTAATTAATTAATTAATTAATTA 1373
|||||
417 oAspSerGlyValTrpValCysSerValAsnThrValAlaGlyMetValG 434
|||||
1374 TGACTCAGGACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1423
|||||
434 luLysProPheAsnIleSerValLysValLeuProLysProLeuAsnAla 450
|||||
1424 AAAGGCTTTCAACATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1473
|||||
451 proAsnValIleAspThrGlyHisAsnPheAlaValIleAsnIleSerSe 467
|||||
1474 CCAATGCTATTGCAATGAGATGAATATATTTGCTATATTAATATATAT 1523

```

```

467 rGluProTyrPheGly 472
|||||
1524 TGAGCCCTTACTTTGGG 1539
|||||
seq_name: /cnu2_6/ptodata/1/ina/5H_COMH.seq:US-08-838-957A-1
seq_documentation_block:
: Sequence 1, Application US/08838957A
: Patent No. 5998187
:
: GENERAL INFORMATION:
: APPLICANT: Breitman, Martin L.
: APPLICANT: Rossant, Janet
: APPLICANT: Dumont, Daniel J.
: APPLICANT: Yamaouchi, Terry P.
: TITLE OF INVENTION: No. 5998187c1 Receptor Tyrosine Kinase
: NUMBER OF SEQUENCES: 32
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Bereskin & Parr
: STREET: 40 King Street West
: CITY: Toronto
: STATE: Ontario
: COUNTRY: Canada
: ZIP: M5H 3Y2
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent in Release #1.0, Version #1.30
: CURRENT APPLICATION DATA: 08/08/98, 957A
: APPLICATION NUMBER: 5998187
: FILING DATE: 23-APR-1997
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Nurdydyk, Linda M.
: REGISTRATION NUMBER: 34,971
: REFERENCE/DOCKET NUMBER: 4153-212
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (416) 364-7311
: TELEFAX: (416) 361-1398
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 4175 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: FRAGMENT TYPE: N-terminal
: ORIGINAL SOURCE:
: ORGANISM: Mus musculus
: STRAIN: CD-1
: DEVELOPMENTAL STAGE: Embryo
: TISSUE TYPE: Heart
: IMMEDIATE SOURCE:
: CLONE: Tek
: POSITION IN GENOME:
: CHROMOSOME/SEGMENT: 4
: MAP POSITION: Between the brown and pmv-23 loci
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 124...3478
:
: US-08-838-957A-1

```

```

alignment_scores:
 Quality: 2407.00 Length: 472
 Ratio: 5.233 Gaps: 0
 Percent Similarity: 97.458 Percent Identity: 91.949
alignment_block:
 US-09-733-764-2_copy_1_472 x US-08-838-957A-1

```

Align seq 1/1 to: US-08-938-957A-1 from: 1 to: 4175

```

1 MetAspSerLeuAlaSerLeuValLeuCysGlyValSerLeuLeuLeu 17
124 ATGGACTCTTTAGCGGCTAGTTCCTCTGTGTGAACTTGTAGCTTGTCTTAA 173
17 rGlyThrValGluGlyAlaMetAspLeuLeuLeuLeuLeuLeuLeu 34
174 TGGAGTATAGAAAGGCGCATGGAATGATGATGATGATGATGATGATGATGAT 223
34 euValSerAspAlaGluThrSerLeuThrCysLeuAlaSerGlyTyrArg 50
224 TGTGTCTCATCGCGAACAATCCCTCCACCTCCCTCCCTCCCTCCCTCCCTCC 273
51 ProHisGluProIleThrIleGlyArgAspPheGluAlaLeuMetAsnGly 67
274 GAAAT 323
67 rHisGlnAspProLeuGluValThrGlnAspValThrArgGluTrpAlaLeu 84
324 GCACCAAGATCCACTGGAGGTATCTCAAGATGTCACCAAGATGTCACCAAGAT 373
84 yLysValValTrpLysArgGluLysAlaSerLysIleAsnGlyAlaTyr 100
374 AAAAAAGTCTTTGCAAGAGAGAAAGCCAGTAAGATTAATGCTGCTTAT 423
101 PheCysGluGlyArgValArgGlyGluAlaIleArgIleArgThrMetLys 117
424 TCTGTGAAGTCTGAGTTCGAGGACAGAGGCTATTAAGGATACGAGACCATGAA 473
117 sMetArgGlnGlnAlaSerPheLeuProAlaIleLeuThrMetThrValAla 134
474 CATGTGTCTAAAT 523
134 sPlysGlyAspAsnValAsnIleSerPheLysLysValLeuIleLysGlu 150
524 ACAGGGCAGATAATCTCAACATACTTTCAAAAGAGGTGTATTAATAAGAA 573
151 GluAspAlaValIleTyrLysAsnGlySerPheIleHisSerValProAr 167
574 GAAGATGAGGATTTACAAAATATGAGGCTCTTATGACACTAGTGGCCGG 623
167 rHisGluValProAspIleLeuGluValHisLeuProHisAlaGlnProG 184
624 CATGAGTATCTGATATTTTAGAAGTTCATTTGCTGCTATCTTATGCTTAT 673
184 IaAspAlaGlyValTyrSerAlaArgTyrIleGlyGlyAsnLeuPheThr 200
674 AGGATGCTGGTGTGCTACTTGGTAACTACATAGCAAGCAAAAGCTGCTTAC 723
201 SerAlaPheThrArgLeuIleValArgCysGluAlaGlnLysTrpGln 217
724 TCAATTTCTATAGTCTATTTTCTGTAATTTCTGTAATTTCTGTAATTTCTG 774
217 yProGluCysAsnHisLeuCysThrAlaCysMetAsnAsnGlyValCysH 234
774 GAGGATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 824
234 sGluAspThrGlyGlyLysIleLysProProGlyPheMetGlyArgThr 250
824 ATGAAGATACCGGGAATGCAATTTGCTGCTTATGCTGCTTATGCTGCTTATG 873
251 CysGluLysAlaValGluLeuHisIleThrPheArgThrCysLysGlnAr 267
874 TCTAGAAAGCTTCTGAGCGGCAACATTTCTGAGGATCTGAGGATCTGAGG 924
267 rCysSerGlyGlnGluGlyCysLysSerTyrValPheCysLeuProAsp 284
924 ATGATGAGGACCAAGGATGCAATTTCTGAGGATCTGAGGATCTGAGGATCTG 974
284 rCysGlyCysSerCysAlaThrGlyTrpLysGlyLeuGluCysAsnGlu 300
974 CTTCAGGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1023

```

```

301 AlaCysHisProGlyPheThrGlyProAspCysLysLeuArgCysSerGly 317
1024 GATGAGGATCTGAGGATCTGAGGATCTGAGGATCTGAGGATCTGAGGATCTG 1073
317 sAsnAsnGlyGluMetCysAspArgPheGlnGlyCysLeuCysSerProG 334
1074 TACCAATCAAGAGATATGATGATGATGATGATGATGATGATGATGATGATGAT 1123
334 LyTrpGlnGlyLeuGlnCysArgGluGlyIleProArgMetThrPro 350
1124 GATGAGGAGGATCTGAGGATCTGAGGATCTGAGGATCTGAGGATCTGAGGAT 1173
351 LysIleValAspLeuProAspHisIleGluValAsnSerGlyLysPheAs 367
1174 CAGATAGAGGATTTGCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 1223
367 nProIleCysLysAlaSerGlyTrpProLeuProThrAsnGluGluMet 384
1224 CCCATCTGCCAAAGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1273
384 hrLeuValLysProAspGlyThrValLeuHisProLysAspPheAsnHis 400
1274 CCTAGTCAAGCCAGATCGGACAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1323
401 ThrAspHisPheSerValAlaIlePheThrIleHisArgIleLeuPro 417
1324 ACAGATGCTTCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTGATCTG 1373
417 oAspSerGlyValTrpValCysSerValAsnThrValAlaGlyMetValG 434
1374 TCATCAAGAGATCTGAGGATCTGAGGATCTGAGGATCTGAGGATCTGAGGAT 1423
434 LuLysProPheAsnIleSerValLysValLeuProLysProLysProLeuAsnAla 450
1424 AAAAGCCCTTCAACAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1473
451 ProAsnValIleAspThrGlyHisAsnPheAlaValIleAsnIleSer 467
1474 GCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1523
467 rGluProThrPheGly 472
1524 TCAGCCCTTACTTTGGG 1539

```

seq.name: /cnp2\_5/plodata/1/ina/5A\_COMP seq.us-08-278-089A-5

seq.documentation\_block:

: Sequence 5, Application US/08278089A

: Patent No. 5681714

: GENERAL INFORMATION:

: APPLICANT: Breitman, Martin L.

: APPLICANT: Rossant, Janet

: APPLICANT: Dumont, Daniel J.

: APPLICANT: Yamaguchi, Terry P.

: TITLE OF INVENTION: No. 5681714e1 Receptor Tyrosine Kinase

: NUMBER OF SEQUENCES: 33

: CORRESPONDENCE ADDRESS:

: ADDRESSEE: Bereskin & Parr

: STREET: 40 King Street West

: CITY: Toronto

: STATE: Ontario

: COUNTRY: Canada

: ZIP: M5H 3Y2

: COMPUTER READABLE FORM:

: MEDIUM TYPE: Floppy disk

: COMPUTER: IBM PC compatible

: OPERATING SYSTEM: PC-DOS/MS-DOS

: SOFTWARE: Patent In Release #1.0, Version #1.30

: CURRENT APPLICATION DATA:

: APPLICATION NUMBER: US/08/278,089A

: FILING DATE: 20-JUL-1994

: CLASSIFICATION: 530





```

874 TGTGAGAAAGTTGTGAGAGGAAATATTTGGAGGAGCTGTAAAGAAAG 923
267 gCysSerGlyGlnGluCysLysSerValThrCysLeuProAsp 284
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
924 GGTAGTGCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 973
284 rGlyGlyCysSerCysAlaThrGlyTrpLeuGlyLeuGlnCysAsnGlu 300
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
974 GTTACGGGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1023
401 AlaCysHisThrGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 317
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
1024 GATGCGCAATCTGGTATATAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1073
317 sAsnAsnGlyGluMetCysAspThrGlyGlyGlyGlyGlyGlyGlyGly 334
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
1074 TACCAATGAAGACATAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1123
334 lyTrpGlnGlyLeuGlnCysGluArgGluGlyIleProArgMetThrPro 350
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
1124 GATGGCAAGCGCTGAGTCTCAGAAAGAGAGAGAGAGAGAGAGAGAGAG 1173
351 LysIleValAspLeuProAspHisIleGluValAsnSerGlyLysPheAs 367
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
1174 CAGATAGAGATTTGTGATATATTAAGTAAGTAAGTAAGTAAGTAAGTA 1223
367 nProIleCysLysAlaSerGlyTrpProLeuProThrAsnGluMet 384
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
1224 CCGCACTCGCAAGCGCTGAGTCTCAGAAAGAGAGAGAGAGAGAGAGAG 1273
384 hrLeuValLysProAspGlyThrValLeuLysProLysAspPheAsnLis 400
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
1274 CCTAGTGAAGCCAGATGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1323
401 ThrAspHisPheSerValAlaIlePheThrIleHisArgIleLeuPro 417
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
1324 ACAGATGTTTTCATTAATTAATTAATTAATTAATTAATTAATTAATTA 1373
417 oAspSerGlyValTrpValCysSerValAsnThrValAlaGlyMetValG 434
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
1374 TGACTCAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1423
434 LuLysProPheAsnIleSerValLysValLeuProLysProLeuAsnAla 450
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
1424 AAAGCGCTTCAACATTCGGGCAAGGCTCTCCAGAGCGCCCTGCACGCC 1473
451 ProAsnValIleAspThrGlyHisAsnProAlaValIleAsnIleSerSo 467
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
1474 CCAATGTCATTCATATGATGATATATATATATATATATATATATATAT 1523
467 rGluProLysPheGly 472
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
1524 TGAGGCTTACTTTGGG 1549

```

```

seq_name: /cqn2_5/ptdata/1/isa/5b_0mb_seq.us 08_220-240A-4

```

```

seq_documentation_block:

```

```

; Sequence 4, Application US/0820240A

```

```

; Patent No. 5955291

```

```

; GENERAL INFORMATION:

```

```

; APPLICANT: Alitalia, Kari

```

```

; APPLICANT: Matkalneo, Marja-Tertuu

```

```

; APPLICANT: Partanen, Juh

```

```

; APPLICANT: Makela, Tomi

```

```

; APPLICANT: Korhonen, Jaana

```

```

; TITLE OF INVENTION: ANTIBODIES RECOGNIZING TIE RECEPTOR

```

```

; TITLE OF INVENTION: TYROSINE KINASE AND USES THEREOF

```

```

; NUMBER OF SEQUENCES: 5

```

```

; CORRESPONDENCE ADDRESS:

```

```

; ADDRESSEE: Marshall, O'Boole, Gersteln, Murray & Borun

```

```

; STREET: 233 South Wacker Drive, 400 South Tower

```

```

; CITY: Chicago

```

```

; STATE: Illinois

```

```

; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.40
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/220,240A
; FILING DATE: 29-MAR-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PC/FT94/00006
; FILING DATE: 08-JAN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/817,800
; FILING DATE: 09-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/167,453
; FILING DATE: 15-DEC-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Gass, David A.
; REGISTRATION NUMBER: 38,153
; REFERENCE/DOCKET NUMBER: 29151/31958
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 474-6300
; TELEFAX: (312) 474-0448
; INFORMATION FOR SEQ ID NO: 4:
; SOURCE CHARACTERISTICS:
; LENGTH: 3845 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 37..3450
; US-08-220-240A-4

```

```

alignment_scores:
Quality: 865.50 Length: 454
Ratio: 2.847 Gaps: 7
Percent Similarity: 66.960 Percent Identity: 37.885

```

```

alignment_block:

```

```

US-09-733-764-2_copy_1_472 x US-08-220-240A-4

```

```

Align: seq 1/1 to: US-08-220-240A-4 from: 1 to: 4845

```

```

14 LeuLeuLeuSerGlyThrValGluGlyAlaMetAspLeuLeuLeuAs 30
||||| ||||||| ||| ||||||| ||||||| ||||||| |||||||
73 CTGCTCTTGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 122
43 nSerLeuProLeuValSerAspAlaGluThrSerLeuThrCysIle 45
||||| ||||||| ||| ||||||| ||||||| ||||||| |||||||
123 CAACCTGCGGCTCAGGACCCGCCCTTCTCTCTGCTGCTGCTGCTGCTG 172
46 ||||||| ||||||| ||| ||||||| ||||||| ||||||| |||||||
173 GATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 216
55 IleThrIleGlyArgAspPheGluAlaLeuMetAsnGlnHisGlnAsp 71
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
217 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 266
71 LeuGluValThrGlnAspValThrArgGluTrpAlaLysLysValVal 88
||||| ||||||| ||| ||||||| ||||||| ||||||| |||||||
267 CTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 315
88 TrpLysArgGluLysAlaSerLysIleAsnGlyAlaLysPheCysGly 104
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
316 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```



```

105 ArgValArgGlyGluAlaIleArgIleArgThrMetLysMetArgGlnG1 121
106 ..
107 ..
108 ..
109 ..
110 ..
111 ..
112 ..
113 ..
114 ..
115 ..
116 ..
117 ..
118 ..
119 ..
120 ..
121 ..
122 ..
123 ..
124 ..
125 ..
126 ..
127 ..
128 ..
129 ..
130 ..
131 ..
132 ..
133 ..
134 ..
135 ..
136 ..
137 ..
138 snValAsnIleSerPheLysLysValLeuIleLysGluGluAspAlaVal 154
139 ..
140 ..
141 ..
142 ..
143 ..
144 ..
145 ..
146 ..
147 ..
148 ..
149 ..
150 ..
151 ..
152 ..
153 ..
154 ..
155 IleThrLysAsnGlySerPheIleHisSerValProArgHisGluValPr 171
156 ..
157 ..
158 ..
159 ..
160 ..
161 ..
162 ..
163 ..
164 ..
165 ..
166 ..
167 ..
168 ..
169 ..
170 ..
171 oAsp..... IleLeuGluValHisLeuProHisAlaClnProGlnA 185
172 ..
173 ..
174 ..
175 ..
176 ..
177 ..
178 ..
179 ..
180 ..
181 ..
182 ..
183 ..
184 ..
185 GGATGGCGGGTTCGCTGGCAAG..... CTCCTCAAAATGTCGACGACCAT 601
186 ..
187 ..
188 ..
189 ..
190 ..
191 ..
192 ..
193 ..
194 ..
195 spAlaGlyValTyrSerAlaArgTyrIleGlyGlyAsnLeuPheThrSer 201
196 ..
197 ..
198 ..
199 ..
200 ..
201 ..
202 ..
203 ..
204 ..
205 ..
206 ..
207 ..
208 ..
209 ..
210 ..
211 ..
212 AlapheThrArgLeuIleValArgArgCysGluAlaGlnLysTyrGlyPr 218
213 ..
214 ..
215 ..
216 ..
217 ..
218 ..
219 ..
220 ..
221 ..
222 ..
223 ..
224 ..
225 ..
226 ..
227 ..
228 ..
229 ..
230 ..
231 ..
232 ..
233 ..
234 ..
235 ..
236 ..
237 ..
238 ..
239 ..
240 ..
241 ..
242 ..
243 ..
244 ..
245 ..
246 ..
247 ..
248 ..
249 ..
250 ..
251 ..
252 ..
253 ..
254 ..
255 ..
256 ..
257 ..
258 ..
259 ..
260 ..
261 ..
262 ..
263 ..
264 ..
265 ..
266 ..
267 ..
268 ..
269 ..
270 ..
271 ..
272 ..
273 ..
274 ..
275 ..
276 ..
277 ..
278 ..
279 ..
280 ..
281 ..
282 ..
283 ..
284 ..
285 ..
286 ..
287 ..
288 ..
289 ..
290 ..
291 ..
292 ..
293 ..
294 ..
295 ..
296 ..
297 ..
298 ..
299 ..
300 ..
301 ..
302 ..
303 ..
304 ..
305 ..
306 ..
307 ..
308 ..
309 ..
310 ..
311 ..
312 ..
313 ..
314 ..
315 ..
316 ..
317 ..
318 ..
319 ..
320 ..
321 ..
322 ..
323 ..
324 ..
325 ..
326 ..
327 ..
328 ..
329 ..
330 ..
331 ..
332 ..
333 ..
334 ..
335 ..
336 ..
337 ..
338 ..
339 ..
340 ..
341 ..
342 ..
343 ..
344 ..
345 ..
346 ..
347 ..
348 ..
349 ..
350 ..
351 ..
352 ..
353 ..
354 ..
355 ..
356 ..
357 ..
358 ..
359 ..
360 ..
361 ..
362 ..
363 ..
364 ..
365 ..
366 ..
367 ..
368 ..
369 ..
370 ..
371 ..
372 ..
373 ..
374 ..
375 ..
376 ..
377 ..
378 ..
379 ..
380 ..
381 ..
382 ..
383 ..
384 ..
385 ..
386 ..
387 ..
388 ..
389 ..
390 ..
391 ..
392 ..
393 ..
394 ..
395 ..
396 ..
397 ..
398 ..
399 ..
400 ..
401 ..
402 ..
403 ..
404 ..
405 ..
406 ..
407 ..
408 ..
409 ..
410 ..
411 ..
412 ..
413 ..
414 ..
415 ..
416 ..
417 ..
418 ..
419 ..
420 ..
421 ..
422 ..
423 ..
424 ..
425 ..
426 ..
427 ..
428 ..
429 ..
430 ..
431 ..
432 ..
433 ..
434 ..
435 ..
436 ..
437 ..
438 ..
439 ..
440 ..
441 ..
442 ..
443 ..
444 ..
445 ..
446 ..
447 ..
448 ..
449 ..
450 ..
451 ..
452 ..
453 ..
454 ..
455 ..
456 ..
457 ..
458 ..
459 ..
460 ..
461 ..
462 ..
463 ..
464 ..
465 ..
466 ..
467 ..
468 ..
469 ..
470 ..
471 ..
472 ..
473 ..
474 ..
475 ..
476 ..
477 ..
478 ..
479 ..
480 ..
481 ..
482 ..
483 ..
484 ..
485 ..
486 ..
487 ..
488 ..
489 ..
490 ..
491 ..
492 ..
493 ..
494 ..
495 ..
496 ..
497 ..
498 ..
499 ..
500 ..
501 ..
502 ..
503 ..
504 ..
505 ..
506 ..
507 ..
508 ..
509 ..
510 ..
511 ..
512 ..
513 ..
514 ..
515 ..
516 ..
517 ..
518 ..
519 ..
520 ..
521 ..
522 ..
523 ..
524 ..
525 ..
526 ..
527 ..
528 ..
529 ..
530 ..
531 ..
532 ..
533 ..
534 ..
535 ..
536 ..
537 ..
538 ..
539 ..
540 ..
541 ..
542 ..
543 ..
544 ..
545 ..
546 ..
547 ..
548 ..
549 ..
550 ..
551 ..
552 ..
553 ..
554 ..
555 ..
556 ..
557 ..
558 ..
559 ..
560 ..
561 ..
562 ..
563 ..
564 ..
565 ..
566 ..
567 ..
568 ..
569 ..
570 ..
571 ..
572 ..
573 ..
574 ..
575 ..
576 ..
577 ..
578 ..
579 ..
580 ..
581 ..
582 ..
583 ..
584 ..
585 ..
586 ..
587 ..
588 ..
589 ..
590 ..
591 ..
592 ..
593 ..
594 ..
595 ..
596 ..
597 ..
598 ..
599 ..
600 ..
601 ..
602 ..
603 ..
604 ..
605 ..
606 ..
607 ..
608 ..
609 ..
610 ..
611 ..
612 ..
613 ..
614 ..
615 ..
616 ..
617 ..
618 ..
619 ..
620 ..
621 ..
622 ..
623 ..
624 ..
625 ..
626 ..
627 ..
628 ..
629 ..
630 ..
631 ..
632 ..
633 ..
634 ..
635 ..
636 ..
637 ..
638 ..
639 ..
640 ..
641 ..
642 ..
643 ..
644 ..
645 ..
646 ..
647 ..
648 ..
649 ..
650 ..
651 ..
652 ..
653 ..
654 ..
655 ..
656 ..
657 ..
658 ..
659 ..
660 ..
661 ..
662 ..
663 ..
664 ..
665 ..
666 ..
667 ..
668 ..
669 ..
670 ..
671 ..
672 ..
673 ..
674 ..
675 ..
676 ..
677 ..
678 ..
679 ..
680 ..
681 ..
682 ..
683 ..
684 ..
685 ..
686 ..
687 ..
688 ..
689 ..
690 ..
691 ..
692 ..
693 ..
694 ..
695 ..
696 ..
697 ..
698 ..
699 ..
700 ..
701 ..
702 ..
703 ..
704 ..
705 ..
706 ..
707 ..
708 ..
709 ..
710 ..
711 ..
712 ..
713 ..
714 ..
715 ..
716 ..
717 ..
718 ..
719 ..
720 ..
721 ..
722 ..
723 ..
724 ..
725 ..
726 ..
727 ..
728 ..
729 ..
730 ..
731 ..
732 ..
733 ..
734 ..
735 ..
736 ..
737 ..
738 ..
739 ..
740 ..
741 ..
742 ..
743 ..
744 ..
745 ..
746 ..
747 ..
748 ..
749 ..
750 ..
751 ..
752 ..
753 ..
754 ..
755 ..
756 ..
757 ..
758 ..
759 ..
760 ..
761 ..
762 ..
763 ..
764 ..
765 ..
766 ..
767 ..
768 ..
769 ..
770 ..
771 ..
772 ..
773 ..
774 ..
775 ..
776 ..
777 ..
778 ..
779 ..
780 ..
781 ..
782 ..
783 ..
784 ..
785 ..
786 ..
787 ..
788 ..
789 ..
790 ..
791 ..
792 ..
793 ..
794 ..
795 ..
796 ..
797 ..
798 ..
799 ..
800 ..
801 ..
802 ..
803 ..
804 ..
805 ..
806 ..
807 ..
808 ..
809 ..
810 ..
811 ..
812 ..
813 ..
814 ..
815 ..
816 ..
817 ..
818 ..
819 ..
820 ..
821 ..
822 ..
823 ..
824 ..
825 ..
826 ..
827 ..
828 ..
829 ..
830 ..
831 ..
832 ..
833 ..
834 ..
835 ..
836 ..
837 ..
838 ..
839 ..
840 ..
841 ..
842 ..
843 ..
844 ..
845 ..
846 ..
847 ..
848 ..
849 ..
850 ..
851 ..
852 ..
853 ..
854 ..
855 ..
856 ..
857 ..
858 ..
859 ..
860 ..
861 ..
862 ..
863 ..
864 ..
865 ..
866 ..
867 ..
868 ..
869 ..
870 ..
871 ..
872 ..
873 ..
874 ..
875 ..
876 ..
877 ..
878 ..
879 ..
880 ..
881 ..
882 ..
883 ..
884 ..
885 ..
886 ..
887 ..
888 ..
889 ..
890 ..
891 ..
892 ..
893 ..
894 ..
895 ..
896 ..
897 ..
898 ..
899 ..
900 ..
901 ..
902 ..
903 ..
904 ..
905 ..
906 ..
907 ..
908 ..
909 ..
910 ..
911 ..
912 ..
913 ..
914 ..
915 ..
916 ..
917 ..
918 ..
919 ..
920 ..
921 ..
922 ..
923 ..
924 ..
925 ..
926 ..
927 ..
928 ..
929 ..
930 ..
931 ..
932 ..
933 ..
934 ..
935 ..
936 ..
937 ..
938 ..
939 ..
940 ..
941 ..
942 ..
943 ..
944 ..
945 ..
946 ..
947 ..
948 ..
949 ..
950 ..
951 ..
952 ..
953 ..
954 ..
955 ..
956 ..
957 ..
958 ..
959 ..
960 ..
961 ..
962 ..
963 ..
964 ..
965 ..
966 ..
967 ..
968 ..
969 ..
970 ..
971 ..
972 ..
973 ..
974 ..
975 ..
976 ..
977 ..
978 ..
979 ..
980 ..
981 ..
982 ..
983 ..
984 ..
985 ..
986 ..
987 ..
988 ..
989 ..
990 ..
991 ..
992 ..
993 ..
994 ..
995 ..
996 ..
997 ..
998 ..
999 ..
1000 ..

```

seq\_name: /rqn2\_6/plotdata/1/ina/5A\_30MB.seq-us-09-188-930-255

```

seq_documentation_block:
; Sequence 255, Application US/09188930A
; Patent No. 6150502
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James Greg
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods For Their Use
; FILE REFERENCE: 11000.1011c1
; CURRENT APPLICATION NUMBER: US/09/188,930A
; CURRENT FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 255
; LENGTH: 1464
; TYPE: DNA
; ORGANISM: Mouse
US-09-188-930-255

```

```

alignment_scores:
Quality: 318.50 Length: 169
Ratio: 3.062 Gaps: 6
Percent Similarity: 61.538 Percent Identity: 38.462
alignment_block:
US-09-733-764-2_COPY_1_472 x US-09-188-930-255

```

```

Align seq 1/1 to: US-09-188-930-255 from: 1 to: 1464
211 CysGluAlaGlnLysIrpGlyProGluCysAsnHisLeuCysThrAlaCy 227
111 ..
62 TGTCACAGGGTGGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCGCTTCG 108
227 smetAsnAsnGlyValCysHisCysAspThrGlyGluCysIleCysPro 244
111 ..
104 CGAGCAAGAGGAGGAGATGTGATGATGATGATGATGATGATGATGATG 158
244 roGlyPheMetGlyArgThrCysGluValAlaCysGluLeuHisThrPhe 260
111 ..
159 CGAGCAAGAGGAGGAGATGTGATGATGATGATGATGATGATGATGATG 208
261 GlyArgThrCysLysGluArgCysSerGlyGluGluGlyCysLysSer 277
111 ..
209 GGCAAGAGGAGGAGATGTGATGATGATGATGATGATGATGATGATGATG 246
277 rValPheCysLeuProAspProThrGlyCysSerCysAlaThrGlyTrp 294
111 ..
247 GGCCCTGTGTATGATGATGATGATGATGATGATGATGATGATGATGATG 296
294 ysGlyLeuGlnCysAsnGluAlaCysHisProGlyPheThrGlyProAsp 310
111 ..

```



```

seq_documentation_block:
: Sequence 1, Application US/08872855
: Patent No. 6:21045
: GENERAL INFORMATION:
: APPLICANT: McCarthy, Sean
: APPLICANT: McCarthy, David
: TITLE OF INVENTION: NOVEL HUMAN DELTA3 COMPOSITIONS AND
: TITLE OF INVENTION: THERAPEUTIC USES THEREFOR
: NUMBER OF SEQUENCES: 23
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: FOLEY, HOAG & ELLIOT LLP
: STREET: One Post Office Square
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02109-2170
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentlib Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/872,855
: FILING DATE: 11 JUN 1997
: CLASSIFICATION: 514
: ATTORNEY/AGENT INFORMATION:
: NAME: Arnold, Beth E.
: REGISTRATION NUMBER: 35,430
: REFERENCE/DOC# NUMBER: MAA-003-02
: TELEPHONE: 617-832-1000
: TELEFAX: 617-832-7660
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2800 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 338..2392
: US-08-872-855-1

alignment_scores:
Quality: 231.50 Length: 248
Ratio: 2.086 Gaps: 14
Percent Similarity: 44.758 Percent Identity: 26.613

alignment_block:
US-08-733-764-2_COPY_1_472 x US-08-872-855-3
Align seq 1/1 to: US-08-872-855-3 from: 1 to: 2055

179 ProHisAlaGlnProGlnAspAla... GlyValTyrSe 190
||||| ||| |||||
391 CCAGAGCCCTTCACACAGATGACACACACACACACACACACACACAGG 440
190 rAlaArgTyrIleGlyGlyAsnLeuPhe... ThrSerAlap 203
||| ||| ||| |||
441 CTCCTAGCTGTGGAGACAACTGTATTATGATATATATATATATATAT 400
203 heThrArgLeu.....IleValIArgArgCysGluAlaGln 214
||||| ||| ||| |||
491 TCACAGCCCTGCGGTACTCTTACCGGTTCATG... TCACAGACAC 534
215 LysTrpGlyProIleCysAsnHisLeuGlySer... 228
535 TACTATGGACACACTCTCTCCGCGCTGTGTAAACAGCGCAATCACCAC 584
226 AlaCysMetAsnAsnGly... 231
585 CGGTACTATGTGTGTATATATATATATATATATATATATATATAT 634
232 ValCysHisGln 236
635 GGACTGGGGAATATGCCAACAGCCTATCTTCTTGGGCTGTCAATGAA 684
236 AspThrGly... GlucyTleCysProProGlyPhe 246
||||| ||| |||||
685 CAGATCGCTACTTCACACACACACACACACACACACACACACACAC 734
246 cMetGlyArgThrCysGluLysAlaCysGluLeuHisThrPheGlyArg 263
||| ||| ||| |||
735 GCAGGCGCGGTGTGTAAAGAA... 756
263 hrCysLysGluArgCysSerGlyGleGlyCysLysSortyrValPhe 279
||| ||| ||| |||
757 TCATCGCCACAAATGCTGTGTG... CACGGCACC 789
280 CysLeuProAspProtyrGlyCysSerCysAlaThrGlyTrpLysGlyLe 296
||| ||| ||| |||
790 TGC AGCACTCGCTGCGAACTTACTCTGTGATCAGCGGCTGTGGGAG 836
296 uClnCysAsnGlu... 300
837 GTTTTGTGACAAATATATATATATATATATATATATATATATATAT 886
301 AlaCys 302
887 ATGGGCGCAACCTGCTCCACAGTGGGTATATATATATATATATATAT 936
303 HisProGlyPheTyrGlyProAspCysLysLeuArgCysSer... 316
||||| ||| ||| |||
937 CGGCCAGGCTACACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 986
317 CysAspAsnGlyGlnMetCys AspArgPheClnGly 328
||||| ||| ||| |||
987 CAGTAAAGGCTGTATATATATATATATATATATATATATATATATAT 1036
329 CysLeuCysSerProGlyGlyGlyGlyGlyGlyGlyGly 341
||||| ||| ||| |||
1037 ACTACTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080

seq_name: /can2_6/prodat3/1/08/872-855-1

```





```

alignment_scores:
 Quality: 209.50 Length: 322
 Ratio: 1.406 Gaps: 20
 Percent Similarity: 46.273 Percent Identity: 24.345

alignment_block:
US-08_743-764-2_COPY_L_472 x US-08-949-366-27 ..
..
Align seq 1/1 to: US-08-949-366-27 from: 1 to: 6677

 211 CYSCLUALALNLYSTFSLPROCHCYSSASHIS..... 222
 |||...|||...|||...|||...|||...
2372 TGGATGACCTGTTCATCTGCAGACAACAAAGCGAGACACAAACTGGTGTGGCTC 2421

 223 LEUDYSTIRALDCYDMEASAGSLVAVCYSHLS...GLUASPTRIG 238

```

[illegible]

3268 CAACCAAGGAAAC 3279

---

100

101

102

103

104

105

106

107

108

109

110

111

112

113

114

115

116

117

118

119

120

121

122

123

124

125

126

127

128

129

130

131

132

133

134

135

136

137

138

139

140

141

142

143

144

145

146

147

148

149

150

151

152

153

154

155

156

157

158

159

160

161

162

163

164

165

166

167

168

169

170

171

172

173

174

175

176

177

178

179

180

181

182

183

184

185

186

187

188

189

190

191

192

193

194

195

196

197

198

199

200

201

202

203

204

205

206

207

208

209

210

211

212

213

214

215

216

217

218

219

220

221

222

223

224

225

226

227

228

229

230

231

232

233

234

235

236

237

238

239

240

241

242

243

244

245

246

247

248

249

250

251

252

253

254

255

256

257

258

259

260

261

262

263

264

265

266

267

268

269

270

271

272

273

274

275

276

277

278

279

280

281

282

283

284

285

286

287

288

289

290

291

292

293

294

295

296

297

298

299

300

301

302

303

304

305

306

307

308

309

310

311

312

313

314

315

316

317

318

319

320

321

322

323

324

325

326

327

328

329

330

331

332

333

334

335

336

337

338

339

340

341

342

343

344

345

346

347

348

349

350

351

352

353

354

355

356

357

358

359

360

361

362

363

364

365

366

367

368

369

370

371

372

373

374

375

376

377

378

379

380

381

382

383

384

385

386

387

388

389

390

391

392

393

394

395

396

397

398

399

400

401

402

403

404

405

406

407

408

409

410

411

412

413

414

415

416

417

418

419

420

421

422

423

424

425

426

427

428

429

430

431

432

433

434

435

436

437

438

439

440

441

442

443

444

445

446

447

448

449

450

451

452

453

454

455

456

457

458

459

460

461

462

463

464

465

466

467

468

469

470

471

472

473

474

475

476

477

478

479

480

481

482

483

484

485

486

487

488

489

490

491

492

493

494

495

496

497

498

499

500

501

502

503

504

505

506

507

508

509

510

511

512

513

514

515

516

517

518

519

520

521

522

523

524

525

526

527

528

529

530

531

532

533

534

535

536

537

538

539

540

541

542

543

544

545

546

547

548

549

550

551

552

553

554

555

556

557

558

559

560

561

562

563

564

565

566

567

568

569

570

571

572

573

574

575

576

577

578

579

580

581

582

583

584

585

586

587

588

589

590

591

592

593

594

595

596

597

598

599

600

601

602

603

604

605

606

607

608

609

610

611

612

613

614

615

616

617

618

619

620

621

622

623

624

625

626

627

628

629

630

631

632

633

634

635

636

637

638

639

640

641

642

643

644

645

646

647

648

649

650

651

652

653

654

655

656

657

658

659

660

661

662

663

664

665

666

667

668

669

670

671

672

673

674

675

676

677

678

679

680

681

682

683

684

685

686

687

688

689

690

691

692

693

694

695

696

697

698

699

700

701

702

703

704

705

706

707

708

709

710

711

712

713

714

715

716

717

718

719

720

721

722

723

724

725

726

727

728

729

730

731

732

733

734

735

736

737

738

739

740

741

742

743

744

745

746

747

748

749

750

751

752

753

754

755

756

757

758

759

760

761

762

763

764

765

766

767

768

769

770

771

772

773

774

775

776

777

778

779

780

781

782

783

784

785

786

787

788

789

790

791

792

793

794

795

796

797

798

799

800

801

802

803

804

805

806

807

808

809

810

811

812

813

814

815

816

817

818

819

820

821

822

823

824

825

826

827

828

829

830

831

832

833

834

835

836

837

838

839

840

841

842

843

844

845

846

847

848

849

850

851

852

853

854

855

856

857

858

859

860

861

862

863

864

865

866

867

868

869

870

871

872

873

874

875

876

877

878

879

880

881

882

883

884

885

886

887

888

889

890

891

892

893

894

895

896

897

898

899

900

901

902

903

904

905

906

907

908

909

910

911

912

913

914

915

916

917

918

919

920

921

922

923

924

925

926

927

928

929

930

931

932

933

934

935

936

937

938

939

940

941

942

943

944

945

946

947

948

949

950

951

952

953

954

955

956

957

958

959

960

961

962

963

964

965

966

967

968

969

970

971

972

973

974

975

976

977

978

979

980

981

982

983

984

985

986

987

988

989

990

991

992

993

994

995

996

997

998

999

1000





```

223 TTTTATATGACCGCTCAAAATGAGGCTGACGCTGCTGCGGCGCG 272
50 qProHisGluProIleThrIleSerArgGluAlaLeuMetAspG 67
273 CCCCCATGAGCCATGACCATATGAGGAGGAGGAGGAGGAGGAGG 322
67 TcHISGInAspProIleuValThrGlnAspValThrArgGluTrpAla 83
423 AGCAGCAGCATGCTGCTGAGGTTATTCAGATCTGCTGCTGAGGAGGCT 372
84 TysLysValValIlePrpLysArgGluLysAlaSerLysIleAsnGlyAlaTy 100
373 AAAAAAGTCTTTTGGCAAGACAGAAAAAGGCTACTAAGATCAATGGTGGCTA 422
100 rPheCysGluGlyArgValArgGlyGluAlaIleAraIleArgThrMetL 117
423 TTTCTGTGAAAGCGGAGTGGAGACAGCAATACCATATGAGGAGGAGGAG 472
117 ySMetArgGlnGlnAlaSerPheLeuProAlaIleThrLeuThrMetThrVal 133
473 AGATGCTGTAAAGGTTCTTTTATCAATATTTTAAATATATATATATAT 522
134 AspLysGlyAspAsnValAsnIleSerPheLysLysValLeuLeuLys 149
523 GACAAGCCACATTAACCTGACATATCTTCCACCAACAGCTATCCATTAAA 572
150 GluGluAspAlaValIleTyrLysAsnGlySerPheIleHisSerValPr 166
573 CAACAACATCCAGGATATTAACAAAAAGGCTGCTTCACTCATTCAGTCC 622
166 oArgHisGluValProAspIleLeuGluValHisLeuProHisAlaGln 182
623 CTGCTATGAAATATATATATATATATATATATATATATATATATATAT 672
183 ProGlnAspAlaGlyValIleValSerAlaArgTyrIleGlyGlyAsnLeuPh 199
673 ACCCAGATGCTGAGGAGTACTGCGGACAGGTAATAGGAGGAGGAGGAGG 722
199 eThrSerAlaPheThrArgLeuIleValArg...ArgCysGluAlaGln 214
723 CACCTCGGCTTCCACCAAGGCTGATAGTCCGAGGAGTATGTCAGAGGAG 772
215 TysTyrGlyProGlyCys AsnHisLeuCysThrAlaLysMetA 229
773 AAGTGTGAGAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 822
229 snAsnGly 231
823 GCAACGCCA 830

seq_name: qb_est1:BB658672
seq_documentation_block:
LOCUS BB658672
DEFINITION BB658672 RIKEN full length enriched, 13 days embryo heart Mus
musculus cDNA clone D330004K12 5', mRNA sequence.
ACCESSION BB658672
VERSION BB658672.1 GI:16492497
KEYWORDS EST
SOURCE house mouse
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 830)
Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A.,
Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Kondo,H., Kondo,
M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,
Okazaki,Y., Okido,T., Saito,R., Sakai,K., Sakai,K., Sano,H., Sasaki,
D., Shibata,K., Shingawa,A., Shiraki,T., Soabe,Y., Suzuki,H.,
Tadami,M., Tagawa,A., Takahashi,F., Takeuchi,Y., Tanaka,T., Taya,T.,
Muramatsu,M., and Hayashizaki,Y.
RIKEN Mouse ESTs (Arakawa,T. et al 2001)

```

JOURNAL  
COMMENT

Unpublished (2001)  
Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Sucho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45 503 9222  
Fax: 81-45-503-9216  
Email: genome-res@qsc.riken.go.jp,  
URL: http://genome.qsc.riken.go.jp/  
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,  
M., Kondo,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full length cDNA libraries for rapid discovery of new  
genes. Genome Res. 10 (10), 1617-1630 (2000)  
waqii,K., Fujiwaki,S., Inoue,K., Toqawa,Y., Izawa,M., Ohara,E.,  
Watanabe,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,  
S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and  
Hayashizaki,Y.

RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multipillar sequencer. Genome Res.  
10 (11), 1757-1771 (2000)

Kondo,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara,  
Y. and Hayashizaki,Y.

Computer based methods for the mouse full length cDNA  
encyclopedia: real-time sequence clustering for construction of a  
non-redundant cDNA library. Genome Res. 11 (2), 261-269 (2001)  
Kondo,S., Shingawa,A., Saito,T., Kiyosawa,H., Yamanaka,I., Aizawa,  
K., Fukuda,S., Hara,A., Itoh,M., Kawai,J., Shibata,K. and  
Hayashizaki,Y.

Computational Analysis of Full-length Mouse cDNAs Compared with  
Human Genome Sequences. Mamm. Genome. 12, 673-677 (2001)  
Please visit our web site (<http://genome.qsc.riken.go.jp>) for  
further details.

## FEATURES

## source

Location/Qualifiers  
1..615  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/clone="D330004K12"  
heart\_lib="RIKEN full-length enriched, 13 days embryo  
heart"  
/tissue\_type="heart"  
/dev\_stage="13 days embryo"  
/lab\_host="DH10B"  
/note "Site 1: Salt; Site 2: BamHI; cDNA library was  
prepared and sequenced in Mouse Genome Encyclopedia  
Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in  
RIKEN, Division of Experimental Animal Research in Riken  
contributed to prepare mouse tissues. 1st strand cDNA was  
primed with a primer [5'  
GAGAGAGAGAGGCGGCAATCGAGTTTTTTTTTTTTTTTTN 3'] cDNA was  
prepared by using trihalose thermo-activated reverse  
transcriptase and subsequently enriched for full-length by  
cap-trapper. Second strand cDNA was prepared with the  
primer adapter of sequence [5'  
GAGAGAGAGATCTCGAGTTAAATTAATTCGCGCGCGCGCGCG 3'] cDNA  
was cleaved with BamHI and XhoI. Vector: a modified  
pBluescript KS(+) after bulk excision from Lambda E1C 1."

BASE COUNT 170 a 128 c 157 g 160 t  
ORIGIN

## alignment\_scores:

Quality: 874.00 Length: 176  
Ratio: 4.994 Gaps: 0  
Percent Similarity: 99.432 Percent Identity: 96.591

## alignment\_block:

MS-09-744-74-2\_copy\_1\_472 x BB658672

Align seq 1/1 to: BB658672 from: 1 to: 615







```

alignment_scores:
 Quality: 656.00 Length: 244
 Ratio: 3.665 Gaps: 3
 Percent Similarity: 73.961 Percent Identity: 45.902

alignment_block:
 US-09-733-764-2_COPY_1_472 x AL547090

 Align seq 1/1 to: AL547090 from: 1 to: 766

122 AAlSerPheLeuProAlaThrLeuThrMetPheValAspLysCysAspAs 138
|||||
29 GCCACCTGCTTCACACAGGTCACACACAGTGTAAACAAAGGAGACAC 78
|||||

138 nValAsnIleSerPheLysLysValLeuIleLysGluGluAspAlaVal 155
|||||
79 CGCTGTACTTTCGACACCTTCACACAGGAGACACACAGACAGCTGACT 128
|||||

155 LeTyLysAsnGlySerPheIleHisSerValProArgHisGluValPro 171
|||||
129 GGAAGACACAGCATCTACTTCTACACCTGACAGGATCAAGCCAG 178
|||||

172 Asp.....IleLeuGluValHisLeuProHisAlaGlnProGlnAs 185
|||||
179 GATGAGGAGTTCCTGTGTGAG.....ATCCCAAAATGTACAGACATC 222
|||||

185 pAlaGlyValTySerAlaArgTyrIleGlyGlyAsnLeuPheThrSerA 202
|||||
223 CACGCCCACTACAGTCCACTTACCTGCAAGTACACACACACACACAG 272
|||||

202 LaPheThrArgLeuIleValArgAcGlyGluAlaGlnLysTrpGlyPro 218
|||||
273 CCTCTTTCGATATATCTGAGAGTCTGAGCTGAGCTGAGCTGAGCTA 322
|||||

219 GluCysAsnHisLeuGlyThrAlaCysMetAsnAsnGlyValCysHisG 235
|||||
423 GGTGTATATTAAGAGTGTATATATATATATATATATATATATATAT 372
|||||

235 AspThrGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyG 262
|||||
473 CCAATGACGGCGAAATGCTATGCCCCCTGCGTACACGGCACCCCTGCG 422
|||||

262 LuLysAlaCysGluLeuHisThrPheGlyArgThrCysLysGluArgCys 268
|||||
423 AACAGAGCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 472
|||||

269 SerGlyGlnGlyCysLysSerLysValPheCysLeuPheAspProGly 285
|||||
473 CCAGGATATACAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 522
|||||

285 pGlyCysSerCysAlaThrGlyTrpLysGlyLeuGlnCysAsnGluAlaG 302
|||||
523 GCGGCG 572
|||||

302 ValHisProGlyPheTyGlyProAspGlyLysLeuArgCysSerCysAsp 318
|||||
573 GTGCGCTGTATATTTTGGAGTGTATTTGGAGTGTATTTGGAGTGTAT 622
|||||

319 AsnGlyClnMetCysAspArgPheGlnGlyCysLeuGlySerProGlyTr 335
|||||
623 AATATGAGAGATTCTATATATATATATATATATATATATATATATAT 672
|||||

335 pGlnGlyLeuGlnCysGluAlaGlnGlyIlePheArgMetThrProLysI 352
|||||
673 GCATGGAGTGCACATCTGACAGAGAGAGAGAGAGAGAGAGAGAGAG 713
|||||

352 LeValAspLeuProAspGlnHisIleGluValAsn 362
|||||
714 TCTTCATCATGGCTGCTGAGAGATGAGAGTTCATAC 745
|||||

seq_name: qb_est2:01743863

Seq documentation_block:
LOCUS R1733863 793 bp mRNA linear EST 20-SEP-2001
DEFINITION 603353149F1 NIH_MGC_94 Mus musculus cDNA clone IMAGE:5360532 5',
mRNA sequence.
ACCESSION R1733863
VERSION R1733863.1 GI:15710876
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: egapbs-remail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11917 row: a column: 13
High quality sequence stop: 548.
Location/Qualifiers
1..793
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:5360532"
/clone_lib="NIH_MGC_94"
/tissue_type="retina"
/lab_host="DH10B (phage-resistant)"
/Note="Organ: eye; Vector: pCMV-Sport6; Site_1: Not 1;
Site_2: SalI; Cloned unidirectionally; oligo-dt primed.
Average insert size 3.3 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 227 a 189 c 220 g 157 t
ORIGIN

alignment_scores:
 Quality: 598.00 Length: 236
 Ratio: 3.398 Gaps: 12
 Percent Similarity: 74.576 Percent Identity: 61.864

alignment_block:
 US-09-733-764-2_COPY_1_472 x B1733863

 Align seq 1/1 to: B1733863 from: 1 to: 793

1 McAspSerLeuAlaSerLeuValLeuCysGlyValSerLeuAlaGlu 17
|||||
124 ATGATATTTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 174
|||||

17 pGlyThrValGlyAlaMetAspLeuIleLeuIleAsuSerLeuPro 33
|||||
174 ATGCATCTACAAAGCCGCAATGACCTGATCTTCACTCAATTCGCAACT 223
|||||

34 LeValSerAspAlaGluThrSerLeuThrCysIleAlaSerGlyTrpAr 50
|||||
224 CTGTGCTGATGCGGAAATCATGCTCATCTGCTGTGGTGGCA 274
|||||

50 gProHisGluPtoIlePheIleCysArgAspPheGluAlaGluMetAsp 67
|||||
274 CACCAAGAGAGAGATATATATATATATATATATATATATATATAT 324
|||||

67 IHisGlnAspProLeuGluValThrGlnAspValThrArgGluTrpAla 84
|||||
324 AGCATCAATGCTGATGATGATGATGATGATGATGATGATGATGATG 474
|||||

```







Align seq 1/1 to AL574074 from 1 to 1011

US-09-733-764-2\_COPY\_11-

Align seq 1/1 to: BF564283 from: 1 to: 371





